Intonation in Luo

Z. S. Ombijah

A thesis submitted for the degree of Doctor of Philosophy in Linguistics

Department of Language and Linguistics

University of Essex

March, 2020

Abstract

This thesis investigates the intonation of Luo. It shows how intonation distinguishes sentences. It also investigates how information structure (i.e. focus, dislocations and topics) is intonationally cued in a sentence. The aim is to establish the phonological and phonetic representation of Luo intonation by examining the factors that contribute to the observed F0 contours. Data were collected in Rorya and Tarime districts in Tanzania, where Luo is predominantly spoken. The materials designed comprise of scripted Luo sentences, Swahili sentences to be translated into Luo and picture-based tasks. The analysis is based on the Auto-segmental metrical theory which maps phonological elements to continuous acoustic parameters (Ladd, 1996, 2008). It is found that Luo is a tone-terracing language with four lexical tones: High, Low, Falling and Rising. The observed downtrends are downstep, declination and final lowering. Downstep is the most significant process, contrasting automatic and non-automatic downstep. The latter has no evidence of floating L and thus attributed to right edge boundary effect. Declination is observed in all-High and all-Low tone sequences, as a phonetic effect. Final lowering is also observed as a final effect in both declaratives and questions. Downstep is also a final effect triggered by a boundary L%. Questions are produced with Pitch Range Expansion triggered by a left edge -H intonational tone. There is no prominence on focused constituents but focus constructions are produced with a higher register. Luo dislocations are asymmetrical, with right dislocations phrased with the main clause while left dislocations are phrased separately from the main clause. Complex clauses, except complementizer clauses, are recursive.

Table of Contents

Acknowledgements

This work could not be a success without financial, academic and moral assistance from various people as well as institutions. First, I would like to thank the Commonwealth Scholarship Commission (CSC) for sponsoring my PhD studies. On top of that I also thank the University of Dar es Salaam for giving me study leave for the entire period that I have been pursuing my studies. This also goes to the University of Essex through the Department of Language and Linguistics and the Proficio programme for allowing me to learn and equip myself with the relevant knowledge that I needed to accomplish this task.

Secondly, I would like to sincerely thank my supervisor, Prof. Nancy Kula, for her tireless academic support throughout my studies. Her invaluable comments and input have contributed to the quality of this work. I would also like to extend my gratitude to Dr. Yuni Kim for taking time to read some of the drafts of this thesis and the entire Phonology research group at the University of Essex for their comments. My special thanks also go to my informants and participants in Tanzania, who took their time to share their knowledge of Luo with me.

Finally, I wish to thank my family and friends for their moral support. Special thanks go to my beloved husband Bless Mwego Maagi for taking care of our little boys Lameck and Jared and keeping the household running smoothly when I was away from home either for private studies, fieldwork or to attend conferences. I also extend my sincere gratitude to our family friends Edson and his wife Sia for giving us moral support and helping us with baby-sitting. To my parents, Saul Ombijah and Sophia Yuda as well my siblings, thank you for believing

in me and encouraging me through your messages and calls. My gratitude also goes to our local church, Praiseway Assemblies of God, at Colchester in the UK and the incumbent Pastors Sean and Joyce McIntyre and former Pastors John and Pauline Holden and all the church members for welcoming us and making us feel at home.

Dedication

To my sons Lameck, Jared and Benjamin

List of Abbreviations

-H	High edge tone
-L	Low edge tone
!H	Downstepped high tone
%T	Initial boundary tone
AdjP	Adjective phrase
AM	Autosegmental-Metrical theory
AP	Accentual phrase
APPL	Applicative
ATR	Advanced Tongue Root
CAUS	Causative
COMP	Complementizer
CONJ	Conjunction
COP	Copula
DEM	Demonstrative
FUT	Future tense
FV	Final Vowel
GEN	Genitive
H	High tone
H%	Boundary high tone
HAB	Habitual tense
HL	High-Low (Falling tone)
HTS	High tone spreading

IMP	Imperative
INF	Infinitive
ip	Intermediate phrase
IP	Intonational phrase
L	Low tone
LH	Low-High (Rising tone)
LOC	Locative
MAE	Mainstream American English
NEG	Negative
NP	Noun phrase
OBJ	Object
OCP	Obligatory Contour Principle
OM	Object marker
PASS	Passive
PhP	Phonological phrase
PL	Plural
POSS	Possessive
PP	Prepositional phrase
PRC	Pitch register compression
PRE	Pitch register expansion
PRF	Perfective aspect
PROG	Progressive aspect
PRS	Present tense

PST	Past tense
Q	Question particle
RECP	Reciprocal
REL	Relativiser
SBJ	Subject
SG	Singular
SM	Subject marker
SVO	Subject Verb Object
T%	Final boundary tone
ToBI	Tone and break indices
TAM	Tense, Aspect, Mood
TBU	Tone bearing unit
V	Verb
VP	Verb phrase
XP	Syntactic phrase
σ	Syllable
φ	Foot
Φ	Phonological phrase
ω	Phonological word

List of Figures

List of Tables

Chapter One

General Introduction

1.0 Introduction

Intonation is a vital part of the grammar of a language, as it signals post-lexical pragmatic meanings such as sentence type (Ladd, 1996). Given that function, it is agreed that all languages, regardless of the nature of their word prosody, have intonation (Bolinger, 1962; Ladd, 1996, 2008; Gussenhoven, 2004). In intonation languages like English, Dutch and German, the intonation of a sentence is determined by pitch accents and edge tones. In tonal languages like Luo, intonation is largely determined by the kind of tones that make up a sentence. This is manifested through the global and local phonetic effects. Recent approaches in the study of intonation involve an examination of the fundamental frequency (F0) which is the physical correlate of pitch from speech waveforms (Pierrehumbert, 1980). In this thesis, an investigation of Luo intonation is given, focusing on how intonation distinguishes sentence types and how information structure is prosodically packaged in a sentence. This involves examining the salient prosodic features used in distinguishing the various syntactic structures. This is useful in determining how well prosodic theories can analyse the interaction between phonology and syntax in Luo.

In this first chapter, general information about Luo, the people and an overview of the grammar are given. An understanding of some of the grammar is required for active engagement in the subsequent chapters, which elaborate on how various aspects of Luo grammar interface with intonation. The organization of this chapter is as follows: Section 1.1 provides some background information about the language and where it is spoken, the people

and the sociolinguistic setup, showing the nature of contact the language has with other neighbouring languages. Section 1.2 describes the important phonological facts about Luo, including the prominent phonological processes. Section 1.3 describes the morphological and syntactic aspects of Luo. The background to the study is provided in section 1.4, the specific research questions in section 1.5 and the scope of the study in section 1.6.

1.1 General background about the language and the people

Luo is a Western Nilotic language belonging to a larger Nilo-Saharan language family. Other languages that belong to the Western Nilotic sub-family are Acholi, Lango, Alur and Padhola. Luo, also known as Dholuo, is spoken mainly in Kenya with a small population in Tanzania. In Kenya the Luo are found in Nyanza province while in Tanzania they are found in Rorya and Tarime districts, in northern Tanzania. According to the *Ethnologue* (Grimes 1996), there are 3.1 million speakers of Luo in Kenya and 223,000 speakers in Tanzania. However, a more recent survey by the Languages of Tanzania Project (2009) reports a smaller number of Luo speakers, at around 170,000, in Tanzania. Historically, the Luo tribe were pastoral nomads who migrated from southern Sudan and entered Kenya through Northern Uganda to the land around Lake Victoria in the late 15th Century. In Kenya they gradually displaced the Bantu speakers, that is, the Kisii and the Luhya (Daigle, 2008). Later on, in the 19th century (around the 1870s) some Luo migrated to Tanzania via Lake Victoria and settled in the Mara region. The second group is believed to have arrived in 1914, after the first world war (personal communication with older speakers in the Luo speaking areas in Tanzania). Through disputes and warfare they managed to displace the Bantu speakers, like the Simbiti, who originally occupied the area.

Rorya district in the Mara region is the main predominantly Luo speaking area in Tanzania. Rorya has four divisions, namely Luo Imbo, Nyancha, Girango and Suba. People occupying the Luo Imbo division claim to be pure Luo. The Rieri, Hacha, and Surwa at the Nyancha division in the northern parts, the Kine at the Girango division in the east and the Simbiti at the Suba division in the south are either Bantu speakers or people who used to speak Bantu languages but shifted to speaking Luo (Kihore, 2005). These people speak Luo as an additional language or have shifted to speaking Luo as their first language. Such a linguistic situation was largely created by intermarriage between the Luo immigrants from Kenya and the indigenous people they found in these areas. Historically, it is claimed that the Luo people offered their women to the men they found in these areas as a strategy to penetrate society and seek acceptance within the community. Mothers being the caregivers who spend most of the time with children taught their children Luo. In Luo Imbo a mixture of languages is not usually observed. This is because it is believed that the Luo fought the indigenous people and conquered their land. On the other hand, Tarime district in the eastern parts is predominantly a Kuria speaking area but with a significant number of Luo speakers especially in the urban areas. People in Rorya as well as in Tarime largely depend on agriculture, livestock keeping, fishing and small industries for a living.

Turning to Luo dialects, there are two officially known dialects: the Trans-Yala dialect spoken in Ugenya, Alego, Imbo and some parts of Gem in Kenya and the Southern Nyanza dialect which is considered the standardized form (Ojal, 2015) spoken in South Nyanza, Siaya and Kisumu in Kenya. The latter is said to extend to Tanzania (Ojal, 2015; Swenson,

2015). Although the Luo variety spoken in Tanzania is said to be the same as the South Nyanza dialect spoken in Kenya there are notable differences especially in the lexicon.

1.2 Phonology

In this section, a short description of the phonology of Luo is provided. The focus is on the sounds (consonants and vowels), the syllable structure and the prominent phonological processes.

1.2.1 Consonants and vowels

Luo has 24 consonants, including the pre-nasalized stops which are considered as phonemic units. The pre-nasalized consonants can appear as onsets or codas of a syllable. The consonants are listed in Table 1.1 below.

Table 1.1 Consonant inventory

Place of articulation / Manner of articulation	bilabial	alveolar	dental	labial	palatal	velar	glottal
plosive	p, b	t, d			c, ɟ	k, g	
fricative		s	θ, ð	f			h
nasal	m	n			ɲ	ŋ	
pre-nasalized	mb	nd				ŋg	
lateral		l					
approximants	w	r			j (y)		

Source: Fieldwork data (2011, 2017, 2018)

Regarding vowels, there are 5 basic vowels, which have an ATR contrast. ATR stands for advanced tongue root. Thus the 5 vowels can be sub-divided into +ATR and –ATR (equivalent to tense and lax vowels, respectively) producing a 10 vowel system (Omondi, 1982; Tucker, 1994). Some recent studies on Luo such as Dimmendaal (2002), Swenson (2010), Ojal (2015), among others, argue for a 9 vowel system. This discrepancy is caused by the disagreement on the status of /a/, and whether it has a [+ATR] counterpart. For those studies assuming 9 vowels, /a/ is considered neutral, with no counterpart, while others like Omondi and Tucker allow [ʌ] as a possible alternant as shown in (1) below.

1. *Vowels*

[+ATR]		[-ATR]	
i	u	ɪ	ʊ
e	o	ɛ	ɔ
(ʌ)		a	

Another aspect to point out regarding Luo vowels is the length. The language does not have a phonemic distinction between long and short vowels. What is noted is the phonetic lengthening of vowels when a word appears utterance-finally, as illustrated in Figures 1.1 and 1.2 below. In Figure 1.1, it is shown that the vowel duration in the word *sûm* 'poison' when in non-final position is 0.094sec, while in Figure 1.2, when the same word is in the final position, the vowel duration is 0.198. As noted earlier, this lengthening does not change the meaning of the word.

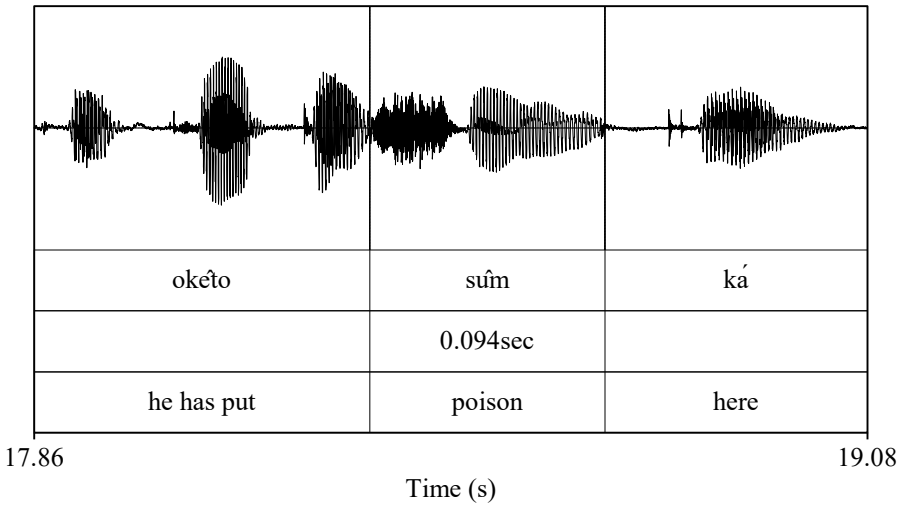


Figure 1.1 The word *sûm* 'poison' in non-final position

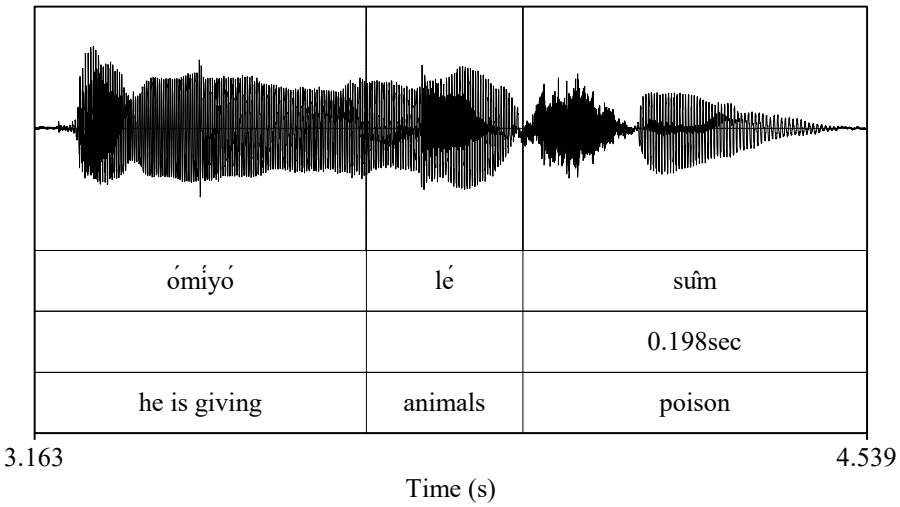


Figure 1.2 The word *sûm* 'poison' in the final position

Another open question regarding Luo vowels is whether the language has diphthongs or not. Maddieson (2013) argues that one can view the movement of vowels from a starting position to a different position as single sounds with an inherent movement (i.e. diphthongs) or view

them as vowels in succession. According to Maddieson these views are based on the researcher's perception. Meanwhile, Batibo (2012) argues that diphthong-like sounds in languages such as Swahili that traditionally have no complex vowels pose descriptive and theoretical problems, the case which is also seen in Luo. Bloomfield (1933) places Luo into class three of secondary phonemes with a sequence of two vowels rather than compound phonemes (i.e. diphthongs) which function as a single unit but Tucker (1994), among many other scholars, considers consecutive vowels in Luo as single units (i.e. diphthongs). As argued by Batibo (2012), this view poses some challenges particularly in analyzing the tonal structure of some words. Consider the Luo example in (2) where the H-tone of the word *pálá* 'ochre' spreads to the first syllable of the word *liet* 'hot' (and not the second). This means if we consider *liet* as monosyllabic, it is expected that the H-tone spreads to both /i/ and /e/ but it only spreads once to the following syllable, which is /i/.

2. pálá + liet → pálá líet
   ochre hot
   'Ochre is hot'

Given such circumstances in analysing the prosodic structure of words, I will consider consecutive vowels as separate sounds that belong to different syllables, as an instance of vowel hiatus.

1.2.2 The Syllable structure

Like most languages, the structure of a Luo syllable consists of an optional onset consonant and a rhyme which is made up of a single vowel. The vowel forms the nucleus of the syllable

which can be followed by a coda. The possible syllable structures and example words are outlined in (3) with the boundary indicated by a dot.

3. *Syllables*

(a) V	e	'in'
	a.o.ra	'river'
(b) CV	pa.la	'knife'
	ki.di	'stone'
(c) VC	ot	'house'
	án	'I'
(d) NCV	mba.ka	'conversation'
(e) CVN	kóm	'chair'
(f) CVC	o.kot	'bell'

All the 24 consonants can occur syllable-initially as onsets. This includes the pre-nasalized consonants as in (3d). Regarding the coda, there are a few restrictions on what can occur at the final position of a syllable. The voiced bilabial stop /b/ and the glottal fricative /h/ cannot occur as codas.

1.2.3 Phonological processes

Luo has several segmental and supra-segmental processes. The latter includes tonal processes which are discussed in chapter 3. In this section, I will highlight the segmental processes which in one way or another can also affect the tonal structure of a word or phrase.

1.2.3.1 Vowel deletion

Vowel deletion is one of the commonly occurring phonological process in Luo. It is a process that is used to resolve hiatus caused by morphological or syntactic concatenation. A language can delete the first or the second vowel depending on its phonological system, but, as Casali (1997) notes, deleting the first vowel is more common cross-linguistically than deleting the second. The examples in (4) demonstrate vowel deletion whereby the first vowel is deleted when followed by another vowel as a result of syntactic concatenation. In this case, the tone of the deleted vowel remains hanging (i.e. floating) and then reattaches to the following syllable. Examples (4a) and (4b) show cases where the initial vowel of the second word is deleted. The same happens to the first vowel of the third word in (4c).

4. (a) /né a-tûo/ → [nátûo]
   PST 1SG-sick
   'I was sick.'

   (b) /má o-cûŋ/ → [mócûŋ]
   REL PRF-stand
   'Who has stood…'

   (c) /wuoro tímó aŋó/ → [wuoro tímáŋô]
   father PROG.do what
   'What is father doing?'

The examples in (5) show vowel deletion resulting from morphological concatenation, in this case plural formation. Strikingly, these examples demonstrate tonal polarity whereby the tone(s) of the singular word is opposite to the tone of the plural word. The tonal polarity observed, however, does not seem to relate to the deleted vowel but to the added vowel (i.e. the plural marker). Yip (2002:159) offers some ways in which polarity can be analysed. One of the proposals is that a H-toned root takes a L-toned suffix, and conversely, a L-toned root takes a H-toned suffix. For Luo, the assumption is that the H or L tone on the suffix spreads left-wards to the root.

5. (a) /cupa + e/ → [cúpé]
   bottle + PL
   'bottles'

   (b) /gowi + e/ → [gópé]
   debt + PL
   'debts'

   (c) /hónó + e/ → [honde]
   miracle + PL
   'miracles'

It is also noted that a vowel can be deleted at the boundary of a lexical word when followed by a CV suffix. In this case, the tone of the deleted vowel does not reattach to the remaining syllables, as shown in (6).

6. (a) /culá + ni/ → [culni]

island + PL

'islands'

(b) /culá + gí / → [culgí]

island + POSS

'their island'

1.2.3.2 Consonant mutation

Most consonants in Luo undergo mutation in certain morphological and syntactic environments. The most observed consonantal changes are approximant hardening and gliding. Odden (2005) defines hardening as a change from a less to a more constricted consonant. This occurs when the word-final consonants /r/, /w/ and /l/ are changed to [c], [p] and [nd] respectively, particularly in plural and possessive cases. Examples are given in (7) below.

7. (a) /bur + e/ → [búcé]

hole + PL

'holes'

(b) /bawo + e/ → [bépé]

timber + PL

'timbers'

(c) /dúól + e/ → [duonde]

voice + PL

'voices'

(d) /dúól ɲaθí/ → [dúónd ɲaθí]

voice child

'child's voice'

(e) /bawo miyo/ → [bap miyo]

timber mother

'mother's timber'

A consonant can change to a glide when another morpheme is added as in (8a) or it can be inserted between vowels, to resolve hiatus as in (8b).

8. (a) /wic + e/ → [wíyé]

head + PL

'heads'

(b) /ó-ði e ot/ → [óði ye ot]

3SG-PROG.go in house

'He is going in the house.'

1.3 Morphology

Luo is a highly isolating language but with some agglutinating features comprising some affixes. The lexicon is mostly characterized by monosyllabic or disyllabic roots. Multisyllabic roots show evidence of being historically inflected. Below is brief analysis of Luo morphology focusing on nominal and verbal affixation.

1.3.1 Nominal morphology

The nominal prefixes *ɟa-*, *ɲa-*, *ra-* and *o-* are attached to verbs or adjectives to form nouns. In most words, when these prefixes are not attached, the roots form bound roots which are meaningless without the prefixes. For that reason the morphological parsing of nominal prefixes in this thesis is ignored. The meaning/function of these prefixes and example words are given in Table 1.2.

Table 1.2 Nominal prefixes

	Prefix Sing/plural	**Meaning/Function**	**Example**	**Gloss**
1	ɟa-/ɟo-	agentive	ɟatêlo/ɟotêlo	leader/leaders
2	ɲa-/ɲi-	diminutive	ɲamin/ɲimine ɲaθí ɲiθîndo	sister/sisters child/children
3	ra-	descriptive	raûm, raŋól	lid, lame person
4	o-	descriptive	oláŋ owát	bell, relative

Source: Fieldwork data (2011, 2017, 2018)

Apart from prefixes, Luo also has two sets of suffixes that can be attached to a noun. These are the plural suffixes *-e, -ni* and *-i*; and the possessive suffixes *-á/-wá, -í/-ú,* and *-é/-gí,* reflecting the first, second and third-person singular/plural, respectively. The possessive suffixes are derived from personal pronouns as illustrated in Table 1.3 below. As already noted above, in case there is a root-final vowel it has to be deleted before attaching any of these suffixes. The plural and the possessive markers can co-occur. In case the co-occurrence creates a vowel hiatus an *–n*, similar to that of an indirect object (see Table 1.3), is inserted as in (9e) and (9f).

9. (a) kóm ‘chair’
   (b) kombe ‘chairs’
   (c) kómbé ‘his chair’
   (d) kombegí ‘their chairs’
   (e) kombené ‘his chairs’
   (f) kombení ‘your chairs

Recall that when a singular noun is H-toned, the plural counterpart is L-toned. For the possessive case, Omondi (1982) observes that the tone of the inflected form is the same before the form was inflected, as can be seen in (9c-f).

1.3.2 Verbal morphology

Luo verbs also undergo affixation. Generally, the verb in its citation form has a final *-o*, which can also be an infinitive marker, as in *tedo* ‘to cook’. Without the final *-o* the mood of the verb becomes imperative, which is manifested by a rising tone. The verb can also be prefixed with pronominal subject markers as well as suffixed with object markers. The

affixes are the short forms of personal pronouns as illustrated in Table 1.3. The full pronouns and the object markers are H-toned while subject markers are toneless and can surface with H or L depending on the associated TAM specifications of the verb (cf. 10a and 10b).

Table 1.3 Verbal prefixes and suffixes

Full pronoun	**Pronominal subject**	**Direct object**	**Indirect object**
án (I)	a- (I)	-á (me)	-ná (for me)
wán (we)	wa- (we)	-wá (us)	-nwá (for us)
ín (you)	i- (you)	-í (you)	-ní (for you)
ún (you)	u- (you)	-ú (you)	-nú (for you)
én (he/she/it)	o- (he/she/it)	-é (him/her/it)	-né (for him/her/it)
gín (they)	gi- (they)	-gí (them)	-nígí (for them)

Source: Fieldwork data (2011, 2017, 2018)

Regarding verbal concatenation in Luo, a pronominal subject marker can co-occur with only one of the object markers as in (10a) and (10b), never with both as in (10c).

10. (a) á-tédó-né

SM1SG-PROG.cook-OBJ

'I am cooking for her.'

(b) a-têd-é

SM1SG-PRF.cook-OM

'I have cooked it.'

(c) *á-téd-é-né

SM1SG-PROG.cook-OM-OM

'I am cooking it for her.'

Apart from the pronominal subject marker, other types of prefixes that can occupy the pre-root slot are the perfective *sé-*, tense marker *né-* and negation *ok-* as in (11) and (12). As usual, the vowel /e/ in *né-* can be deleted to resolve hiatus as discussed in section 1.2.3 and its tone carried by the following syllable. In Luo, however, the past tense marker *né-* and negation *ok-* can occur as independent words. Note that *né-* is derived from the independent adverb *néné* 'long ago'. We return to tense marking in section 1.3.3 below.

11. né-o-sé-têdo → [nósétêdo]

PST- SM3SG-already-PRF.cook

'He had already cooked.'

12. né-ok-o-têdo → [nókotêdo]

PST-NEG- SM3SG-PRF.cook

'He did not cook.'

Other than the infinitive and object markers, the habitual tense marker *-ga*, imperative marker *-í* (for singular) and *-urú* (for plural) and the reciprocal *-r-* can also occupy the post root/stem position as shown in (13a-d), respectively.

13. (a) á-tédó-ga

SM3SG-cook-HAB

'I usually cook.'

(b) ted-í

cook-IMP

'(You) cook.'

(c) ted-urú

cook- IMP

'(All of you) cook.'

(d) ó-tédó-ré

SM3SG-cook-RECP

'It is cooking itself.'

1.3.3 Tense marking

In the literature on Luo grammar there is a disagreement on whether tense marking in the language is morphological or syntactic. As we have seen in examples (11), (12) and (13a) above, the morphemes *né-* and *-ga* are attached to the verb to show past and habitual tenses,

respectively. Although this has become a tradition in the analysis of Luo grammar, it has been criticized especially in the early studies of the language. Omondi (1982) argues that Luo does not inflect tense because there is no form that can be said to realize a past action. What we see are reduced forms of time adverbials which are mistaken for prefixes. She adds that the reduced forms can stand unprefixed to the verb. Her list of time adverbials and their corresponding reduced forms is given below.

14.	**Time adverbials**	**Reduced forms**	**Gloss**
a)	nénde	né	'early today'
b)	nyóro	nyó	'yesterday'
c)	nyoca	nyoc	'day before yesterday'
d)	néné	né	'sometime in the past'
e)	yandé	yand	'a few days ago/recently'
f)	kîny	-	'tomorrow'
g)	âng'	-	'later today'

In support of the recent tradition of morphological tense marking, Dimmendaal (1995:36) argues that so long as these adverbs may precede verbs as independent sentence level markers in all Nilotic languages, they can also develop structurally into preverbal clitics (proclitics) or affixes which are incorporated into the verb to mark tense. Dimmendaal, however, points out that this kind of synchronic change is due to massive interference from the neighbouring Bantu languages.

In my study, I noticed that speakers frequently use the reduced form *né* to indicate past action than the full form time adverbial. It can also be seen in (14f) and (14g) that the adverbs *kîny*

'tomorrow' and *âng'* 'later today' are not reduced or contracted. Thus, future tense is marked by the word *bíró* which means 'to come'. The word *bíró* can be reduced to *bo* as shown in (15), in this case, both *bíró* and *kîny* co-occur as tense markers.

15. ɲiθîndo, bíró tugo e kúóyó kîny godiocíèŋ
children come play in sand tomorrow afternoon
'Children will play in the sand tomorrow afternoon.'

In this thesis, I consider both morphological and syntactic methods as appropriate and acceptable ways of marking tense in Luo, as the choice of whether to use morphemes or adverbs entirely depends on the speaker and the context in which the language is used. In casual or informal speech the former is preferred than the latter.

Regarding aspect, there are two types identified: imperfective and perfective aspects. The former indicates actions that are still taking place while the latter indicates actions that are completed. In most cases, aspect in Luo is phonologically marked by changing the tonal structure of the verb. This is explained in detail in chapter 3 section 3.1.3 under grammatical tones. In addition to the special tonal structure, the perfective aspect can also be morphologically marked by a prefix *o-*. This aspectual prefix is homophonous to the third person singular prefix in (13d) but with a different syntactic role as illustrated in (16).

16. (a) Jane o-têdo réc
Jane PRF-cook fish
'Jane has cooked fish.'

(b) mon o-têdo réc

women PRF-cook fish

'The women have cooked fish.'

1.4 Background to the study

Generally, Luo is a well-studied language in almost all the grammatical aspects. Some of these aspects are pointed out in the previous sections. Concerning Luo intonation structure, however, less has been said. There have been a few studies that shed some light on Luo tonology such as Tucker and Creider (1975), Creider (1978), Omondi (1982) and Tucker (1994). Tucker and Creider (1975), as well as Tucker (1994), generally argue that downtrends such as downstep and downdrift are the prominent intonation features found in Luo sentences. Their studies are generic as they do not systematically investigate different types of constructions such as comparing statements and questions or focused and non-focused constructions. Creider (1978) goes further by giving three characteristics of Luo intonation and analysing the phonetic correlates of the so-called tone groups. He first identifies three types of tone groups namely declarative, interrogative and sustaining. He argues that full word stress is found in the last word in a tone group. Such an intonation contour is found in intonation languages like English and German, which have pitch accents as tonal events at particular points in a phrase or utterance. For tonal languages like Luo, it is generally observed that lexical tones locally interact at various points in a sentence and do not necessarily create a plateau until the end of an utterance where an intontional tone is manifested. Creider, nevertheless gives an overview of Luo intonation and sentence type that both declaratives and interrogatives are perceived with a gradual fall in pitch, although the latter type is produced at a higher pitch range. Contrary to Creider, Omondi (1982) claims

that questions in Luo end in a rising contour. Her study is not specific to the intonation of Luo, as it analyses all the major syntactic structures of the language. Although their studies are impressionistic, their arguments are used as a starting point to investigate further the intonation of Luo declaratives and questions and also include other types of syntactic constructions such as focus, topics and dislocations.

Furthermore, the instruments and approaches that the former scholars used cannot determine clearly how intonation is phonologically and phonetically modelled in Luo. This means that there is no detailed work on Luo intonation which is based on current instrumental approaches that model the phonological and phonetic representations of intonation. Therefore, in this thesis, Luo intonation structure will be analysed based on the Auto-segmental metrical theory, which describes both phonological and phonetic accounts of intonation (Ladd, 1996; 2008). This involves investigating the fundamental frequency (F0) contours from computed waveforms, instead of relying on musical instruments.

1.5 The research questions

This study aims to establish the phonological and phonetic representation of Luo intonation by examining the factors that contribute to the observed F0 contours. Thus the following questions are addressed:

(a) How many tones does Luo have? Which tones are phonologically contrastive and which ones are phonetically manifested?

(b) What are the phonetic correlates of declarative sentences?

(c) How are the Luo declarative sentences prosodically different from questions?

(d) Are focus, topics and dislocations prosodically cued in Luo? If yes, what are the cues?

(e) How are complex sentences prosodically cued? Are they recursive?

1.6 Scope

All previous work on Luo tonology were conducted on the Kenyan variety. Regardless of the claim that the Luo spoken in Kenya is the same as that spoken in Tanzania (Ojal, 2015; Swenson, 2015), there is a prevailing notion among speakers that Tanzanian Luo is simplified. In this thesis and at this stage I do not intend to compare the two varieties. For convenience, the data in this thesis are based on the Luo variety spoken in Tanzania. Other varieties are not covered. It should be noted that the Nyanza dialect spoken in Kenya is the one which is said to extend to Tanzania. Therefore, the findings from this research may apply to the other varieties of Luo as well.

Regarding the scope of the data used in this study, most elicitations were based on read text with only a few elicitations coming from spontaneous speech. Given that the latter was not systematically tested, we can only assume that the findings here represent some of what would be attested in natural speech. Furthermore, the intonation analysis herein only focuses on simple declarative sentences, complex sentences, questions and information structure. Only two types of questions, namely, yes-no (polar) questions and content questions are analysed. The information structure investigation includes contrastive and non-contrastive focus, dislocations and topics. Other types of questions and other information structure constructions will have to be considered in future work.

Chapter Two

Literature Review and Theoretical Background

2.0 Introduction

This chapter reviews the intonation and prosodic structures of various languages based on their prosodic typology. The chapter covers a sample of languages identified to have tone, stress or lexical pitch accent at the word level. The main focus is on how tone, stress and accent, as supra-segmental or prosodic features, interact with intonation across languages. As defined by Hyman (2007) tone is pitch specified at word level to dustinguish lexical meanings or grammatical functions. Stress is also a metrical feature identified at the word level but does not necessarily distinguish lexical meanings. According to Hyman (2007) accent combines properties from tone and stress prototypes, which in turn, produces a mixed, ambiguous and intermediate system. Hyman thus concludes that there is no pitch-accent prototype. On the other hand, Gussenhoven (2004) argues that pitch-accent is simply the location where intonational tones are placed. Gussenhoven (2004:12) adds that all languages have phonological stress in the sense of foot structure. While other languages have observable phonetic salience for stress, other languages have none. Although tone, stress and accent have different functions, phonetically, they all have similar pitch correlates. In this chapter, we will discuss the salient features of intonation in languages with stress, lexical tone and accent.

The chapter is organized as follows. Section 2.2 explains what intonation is. Section 2.3 presents the current models used in describing the phonology of intonation. Section 2.4 shows the differences and similarities in modelling intonation in non-tonal and tonal

languages. Section 2.5 discusses the phonological representations as well as the phonetic implementations of different types of sentences in both non-tonal and tonal languages. Section 2.6 briefly talks about prosodic structures and how they are phonetically cued.

2.1 What is intonation?

The notions stress, tone, accent and intonation are all supra-segmental aspects but with different implcations (Gussenhoven, 2004). As already introduced above, the first three are identified at the word-level. Intonation, on the other hand, is post-lexical. As Hyman and Monaka (2011) put it, intonational features should not be the same features present earlier in the phonology of the word. They should rather be in the syntax. Thus, regardless of whether a language has lexical prosodic features or not they all have intonation structure (Bolinger, 1962; Ladd, 1996, 2008; Gussenhoven, 2004). Ladd (1996, 2008) defines intonation as the use of supra-segmental phonetic features to convey post-lexical pragmatic meanings in a linguistically structured way**.** Similarly, Gussenhoven (2004:12) adds that it is the use of phonological tone for non-lexical purposes such as expressing phrasal structures and discourse meaning.

Ladd's (1996) and Gussenhoven's (2004) definitions are based on the instrumental approach. Intonation approaches will be discussed shortly in section 2.2. The motive behind the instrumental approach is to evaluate how speech is perceived in relation to its phonetic cues. Ladd (1996) adds that it is important for supra-segmental features to have meanings at post-lexical or sentence level. An example would be the use of supra-segmental features to distinguish statements from questions or focus from non-focused information. Secondly,

these supra-segmental features should be linguistically structured in terms of being categorically distinct entities. This part excludes other paralinguistic features such as tempo or loudness which signal the speaker's state at the time of speaking. For a linguistically structured feature an example would be a low ending indicating a statement and a high ending a question.

Another approach to intonation is based on the impressionistic view. This approach is basically auditory and was preferred by teachers for practical purposes of teaching a foreign language. O'Connor and Arnold (1973) is one of the earliest works on intonation of English based on the impressionistic approach. To them, intonation should be significant, systematic and characteristic. It is significant in the sense that utterances which differ in their intonation yield different meanings. Secondly, these different tunes are not randomly produced but are rule-based and systematic in their usage and therefore they should be learned. And lastly, intonation is characteristic because the pitch patterns or tunes are language specific.

2.1.1 Universal intonational meaning

Gussenhoven (2002, 2004) gives an acoount of the universal meaning of intonation based on the universal tendencies in the use and interpretation of F0. These tendencies are explained under the so-called biological codes. Gussenhoven's idea of biological codes in pitch variation is based on the speech-production mechanism. First, the rate of vibration of the vocal cords depends on the size of the larynx. The smaller the larynx the faster the vibration. Secondly, air pressure is generated through phases which in turn form breath groups. Lastly, the speech production process may be executed with lesser or greater precision

(Gussenhoven, 2004:79). Thus, the biological codes include frequency code, effort code and production code. Gussenhoven argues further that apart from intonational meanings being universal they can also be language specific. The universal part is exercised in the phonetic implementation as stipulated by the speech-production mechanism while language specific meaning is in the intonational morphology and phonology. It should be noted that the term universal here, as used by Gussenhoven, refers to a form-meaning relation in the phonetic implementation and not a generic view of universal structural elements.

2.1.1.1 The frequency code

As introduced above, frequency code refers to the rate of vibration of the vocal cords. A smaller larynx has smaller and lighter vocal cords, which tend to vibrate faster producing higher pitches (Gussenhoven, 2002; 2004). Gussenhoven stipulates that the frequency code can have other affective or informational interpretation. Affective interpretation can be things like higher pitch denoting friendliness, politeness, vulnerability as opposed to aggression, scathing and confidence for lower pitch. Regarding informational interpretation, higher pitch has been attributed to uncertainty and lower pitch to certainty, hence questions versus assertions. Such informational interpretation has been grammaticalized in many languages. Gussenhoven reports that over 70% of the world's languages have a rising pattern for questions and a falling contour for statements.

2.1.1.2 The effort code

The effort code, on the other hand, refers to the energy expended on speech production. Here, the more the energy used in production of speech, the greater the articulatory precision and the wider excursion of pitch movements. Exploitation of effort code makes use of pitch-span

variation to signal different meanings. An example of informational interpretation from the effort code is when a speaker uses a wider pitch-span for emphasis such as focus. Thus, focus is grammaticalization of the effort code. Languages have varied ways in which focus is marked. Gussenhoven (2004) reports that in most languages focused information has a wider pitch excursion, which means intonation plays a vital role in its determination. Another grammaticalized construction in this respect is negation. Here the speaker reduces the excursion size rather than increasing it. At this juncture the highs and the lows are brought together implying that the speaker is withdrawing information. Concerning affective interpretation, wide excursions indicate being authoritative, insistent and enthusiastic while narrow excursions indicate lack of commitment and being uninterested (Gussenhoven, 2004:88).

2.1.1.3 The production code

As it was for the effort code, the production code also involves energy. In the production code, however, the process of energy generation is in the form of breath groups or phases. The air pressure is higher at the beginning of a breath group and lower towards the end causing a declining trend. According to Gussenhoven, the significance of declination is not on the slope but on the utterance edges, that is, high vs low edges. He adds that normally high beginnings signal new topics while low beginnings signal the continuation of a topic. Conversely, high endings signal continuation while low endings signal finality (Gussenhoven, 2004:89).

To conclude on the biological codes, Gussenhoven argues that despite being universal they are conventionalized within specific languages. Different speech communities may interpret

similar phonetic implementation differently or extensively use a certain code rather than another. In this thesis, I adopt Ladd's (1996, 2008) definition of intonation by showing how supra-segmental features yield post-lexical meanings which are linguistically structured and, at the same time, pointing out the language specific intonational tendencies in Luo as proposed by Gussenhoven (2002, 2004). In the following section different intonation models are discussed.

2.2 Intonation models

Intonation is an important part of a language's grammar and thus it needs a model that would capture useful generalizations about the system; what is contrastive in the system and what would be predicted from the system, with the intention to make a speaker or a learner use appropriate melody in a given situation (Arvaniti, 2016). Therefore, in order to systematize intonation, different models have been developed.

Until the late 1970s there were only two approaches to studying intonation. These were the impressionistic and instrumental approaches (Ladd, 1996). The impressionistic approach simply looks at how tunes are heard, interpreted and understood. It was preferred by linguists and language teachers interested in the description of intonation for practical reasons. It was meant to improve the pronunciation of foreign language speakers, particularly learners of English. This approach was interested in categorically distinct elements such as pitch phonemes and nuclear tones (Ladd 1996). An example of intonational works based on this approach is that of Palmer (1922) and O'Connor and Arnold (1973). The Instrumental approach, on the other hand, equates intonation to F0 contours. This approach uses acoustic

cues to evaluate intonational phenomena such as syntactic or pragmatic notions like finality, continuation and interrogation; emotional states like anger, surprise or boredom; as well as word and sentence stress (Ladd, 1996:12). It was preferred by experimental psychologists and phoneticians who were interested in speech perception and the associated phonetic cues. The instrumental approach is preferred in current studies of intonation as we will see in the discussion below. It should be noted that the survey of the intonation models in this chapter focuses on linear models, excluding superpositional models such as Fujisaki's model.

2.2.1 British traditional model

Palmer's (1922) and O'Connor and Arnold's (973) models are based on the British tradition, with the intention to describe the intonational patterns of British English. The British tradition is mainly impressionistic. The main objective of Palmer (1922) and O'Connor and Arnold (1973) was to describe the endings of English utterances. Palmer (1922) points out that English intonation contours comprise the nucleus, the head and the tail. The nucleus is the obligatory part which contains the most prominent stress. That stress is normally on the last word in a multi syllabic utterance. The head is the stretch of contour which precedes the nucleus, while the tail is the stretch of contour following the nucleus. Palmer identifies four tone groups labelled as Tone-group 1, Tone-group 2, Tone-group 3 and Tone-group 4. The tones represented in each group are falling, high-rising, falling-rising and low-rising, respectively. The identified tones are said to have specific semantic functions. For example, tone group 1, with a falling contour on the nuclear word can signal statements, special questions, commands, general questions and words in isolation. Tone-group 2, with a high-rising tone can signal echo questions or statements. Tone-group 3 which contains a falling-

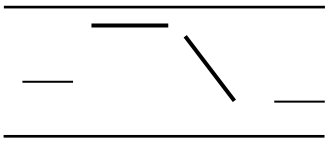
rising contour can signal statements or commands. Tone group 4, with a low-rising contour, signals parting greetings or statements with reassuring character (Palmer, 1922:86).

Palmer's intonation model is reflected further in the work of O'Connor and Arnold (1973) but with some modifications. O'Connor and Arnold's model has four parts instead of three. These parts are pre-head, head, nucleus and tail. According to O'Connor and Arnold, a 'pre-head' is a syllable(s) occurring before the stressed syllable of the first accented word while the 'head' is that part of the word with a stressed syllable (the first accented word but not the nucleus). A pre-head contains a low or high tone while the head contains a low, high, falling or rising tone. The 'nucleus', like in Palmer's (1922) definition, is the stressed syllable of the last accented word in a tone group. The nuclear tone is located in the nucleus. The last part in a contour is the 'tail'. The tail comprises a string of syllables (stressed or unstressed) which follow the nucleus.

Contrary to Palmer (1922), O'Connor and Arnold (1973) came up with six nuclear tones and ten tone groups. The nuclear tones are Low Fall, High Fall, Rise Fall, Low Rise, High Rise and Mid-Level. These are like what Palmer (1922) refers to as tone groups. To O'Connor and Arnold 'a tone group' reflects affective meaning (showing attitude of the speaker) and is also in conjunction with the five sentence types, namely statements, wh-questions, yes-no questions, commands and interjections. The possible semantic functions include completeness or definiteness, seriousness, urgency, calmness, friendliness, liveliness, acceptance, etc. (O'Connor and Arnold, 1973). O'Connor and Arnold represent the contour

pattern in a musical stave-like annotation. An example of a sentence with all the four parts is given below in (17).

17. *You'd be surprised!* (Low Pre-head +) High Head + High Fall (+Tail)

Tune 

O'Connor and Arnold (1973:133)

2.2.2 Pierrehumbert model

One of the most renowned models of intonation based on the instrumental approach is that of Pierrehumbert (1980). The 1980 work is the original version of Pierrehumbert's description of English intonation. Other modified versions are those developed with Mary Beckman in 1980 and 1986. Their work aimed at developing the phonological and phonetic representations of English intonation by identifying possible underlying representations and then determining their phonetic properties (i.e. the surface representation). Pierrehumbert's work remains the most influential in current times in the phonology of intonation, as it has been used as a foundation in the development of the ToBI model and Autosegmental-Metrical theory.

Pierrehumbert (1980) proposes that intonation has three phonological characterizations: a) There are allowable phrasal tunes generated by a sequence of L and H tones; b) These phrasal tunes have metrical representations determined by the patterns of stressed and unstressed syllables; c) There are rules which align tune with text. These phonological

characterizations then give rise to three important parts of a pitch contour: pitch accent, phrasal accent and boundary tone.

Pitch accent is the prominence given to a syllable to portray the intended meaning. It can occur as a single tone (i.e. monotonal) or as a pair of tones (i.e. bitonal). Note that pitch accents are marked by an asterick (*) to indicate an accented syllable. In bitonal contexts, there is a leading or trailing tone which precedes or follows a starred tone marked by a raised hyphen. The use of hyphens was later discarded in the subsequent works on intonation. The possible pitch accents are as follows: H*, L*, L+H*, L*+H, H+L*, H*+L. Pierrehumbert's 1980 version recognizes a seventh pitch accent, H*H, which is eliminated in the later modified versions.

Apart from pitch accents, a contour is also composed of phrase accents and boundary tones, which together form what she calls edge-tones. Pierrehumbert identifies two phrase accents, H^- and L^- and two boundary tones H% and L%. Phrase accents are found at the right edge of an intermediate phrase (ip) and boundary tones are found at the right edge of an intonation phrase (IP). Generally, Pierrehumbert notes that a low boundary (L%) indicates certainty or completeness and is found at the end of statements or calls. On the other hand, high boundary (H%) indicates uncertainty, a typical contour for questions or incompleteness. This model is schematized below:

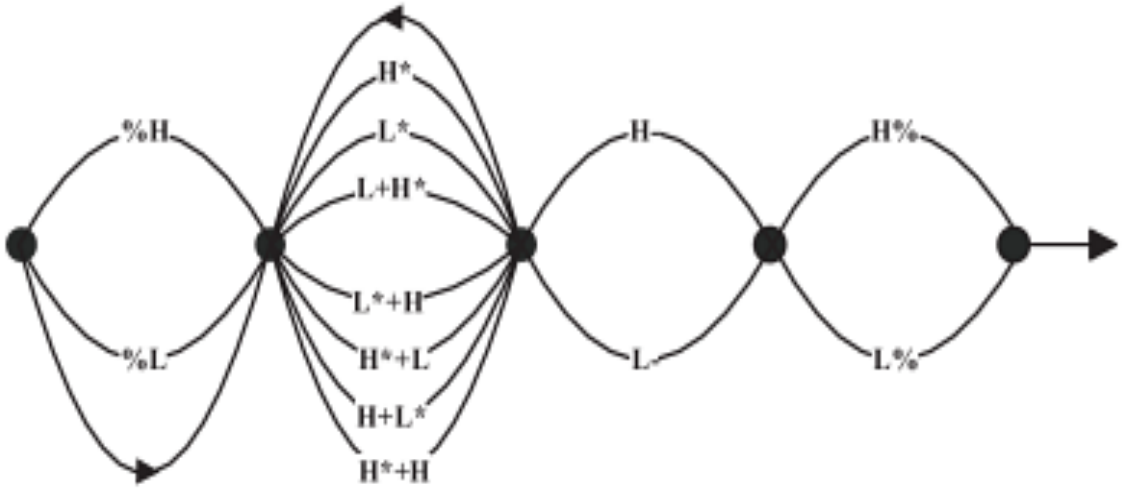


Figure 2.1 Pierrehumbert model showing pith accents, phrase accents and boundary tones

Illustrating how the model works, Pierrehumbert (1980) gives an example of the word 'Anna' which is stressed on the first syllable and unstressed on the second yet yielding different melodies by combining pitch accents, phrase accents and boundary tones. Consider the following contour structures.

18.

(a) H* L- L%

(b) H* L- H%

(c) H*+L- H- L%

(d) L*+H- L- H%

(e) L* H- H%

Focusing on the accented syllable (marked by an asterick), in (a) we see a peak at the beginning followed by a fall to the bottom of the speaker's pitch range, a melody Pierrehumbert (1980) identifies as an answer to a question. (b) is like (a) except that there is a rise at the end, and is an incomplete answer to a question. (c) begins with a complex melody (i.e. H*+L) and then stops far short of the bottom of the speaker's pitch range. Such a melody can function as a call out to Anna. In (d) the contour starts lower and maintains this level before rising, as an indication of incredulousness. And lastly, (e) is very low on the accented syllable with a rise at the end, a common pattern for question intonation. A similar kind of melodic patterning can be observed when words are put into phrases or utterances. However, in such larger constituents the peaks and valleys seen in F0 contour will depend on which word carries the nuclear stress of the phrase or utterance. Pierrehumbert notes that the peaks and valleys portray the important tonal events which occur as falls or rises at the specified points in a phrase or utterance. What occurs between the tonal events is a transition referred to as F0 plateauing.

Although pierrehumbert's approach has been widely accepted and extended in other languages, it has also been criticized. According to Ladd (2008), there has been a disagreement on the difference between H* and L*+H. He elaborates that if the two are followed by L phrase accents, they bring about a falling contour represented as H*L and rising-falling contour represented as L*+HL. Thus, some scholars consider L*+HL as a variant of falling with their difference based on paralinguistic features but Pierrehumbert considers both contours as distinct linguistic categories.

Another criticism also pointed out by Ladd (2008) is on the distinction between a bitonal L+H* and a monotonal H*. This is because L+H* is characterized by pitch movement from the preceding syllable rather than the following syllable. In that sense, it conflicts with H* which also indicates what Ladd (2008) calls 'pitch jump' from a preceding syllable. He notes that the two are difficult to distinguish phrase-initially on an accented syllable because they both suggest a rising contour but one with a L-tone, while the other has no tone.

2.2.3 ToBI model

ToBI which stands for Tone and Break Indices is a standard convention for labeling prosody. It was first developed in a series of four workshops conducted between 1991 and 1994 which included experts from various disciplines such as engineering, psychology, computer science, phonetics and others (Beckman et al. 2005). ToBI is not a theory in itself but a convention for annotating prosody. The original system was meant to describe the intonation of American English thus called Mainstream American English ToBI (MAE_ToBI) but was later applied to other dialects of English and to other languages as well. Therefore, the name MAE_ToBI is reserved for the original system while ToBI is generally used for other developmental works.

ToBI convention has six obligatory parts (see Figure 2.2) divided into two continuous phonetic records and four symbol strings (Beckman et al. 2005). The former includes an audio recording of the utterance graphically represented by the waveforms and the fundamental frequency (F0) contour. The symbol strings have four tiers for labeling tones, words, break indices and miscellaneous events. These are further described as follows.

The Tonal tier, which follows Pierrehumbert's analysis, shows that the tune of an utterance is transcribed as a linear sequence of pitch events that are sparsely distributed across the text. In the ToBI model, as well as in Pierrehumbert's model, a pitch contour has three basic tones including phrase accents (H- or L-); boundary tones (H% or L%); and pitch accents (L*, H*, L+H*, L*+H, H+!H*). Contrary to the Pierrehumbert model, which had seven pitch accents in the original 1980 analysis, and six in later developments, ToBI has five pitch accents whereby the H*+L accent is not used. Moreover, ToBI adopted other labels such as downstepped (!H) to mark a compressed pitch range and (X*?, X-?, X%?) to mark uncertainty (Silverman et al, 1992; Beckman et al, 2005).

The second tier is the word tier. This is simply an orthographic transcription of an audio recording. After this follows another important tier called the Break Indices. The Break Index tier accounts for the prosodic phrase boundaries which are indicated by numbers depending on the boundary strength. As summarized in Beckman et al (2005) there are four basic break index values scaled from 0 to 4. 0 shows a very close inter-word juncture; 1 shows ordinary phrase-internal word end junctures; 3 shows intermediate phrase end marked by a phrase accent; and 4 shows intonational phrase end marked by a boundary tone. Break index 2 is left out in this scaling because it shows a mismatch between tones and breaks by marking ambiguous inter-word junctures inappropriate for 1 and 3. Lastly, the miscellaneous tier labels other prosodic events such as hesitations, disfluencies, breaths, laughs, false starts and restarts (Silverman et al, 1992). Silverman points out that these events have less importance

in the prosodic analyses and are very rare in laboratory recordings, but it is still a pervasive problem in real-world applications.

The most important tiers are the tonal tier and the break indices which give the system its name. Later analyses which refer to ToBI transcription, show interest in the tonal tier only, ignoring the break indices (Ladd, 2008). As for the diacritics, ToBI uses similar diacritics as those in Pierrehumbert model and in Autosegmental-Metrical theory (see below), except for a minus (-), as opposed to a raised hyphen used in Pierrehumbert model, to indicate phrase accents. ToBI model is illustrated further in Figure 2.2 below. In this figure, the utterance 'Will you have marmalade, or jam?' with L* H- L* H-H% contour, shows a final rise as an illustration of a yes-no question.

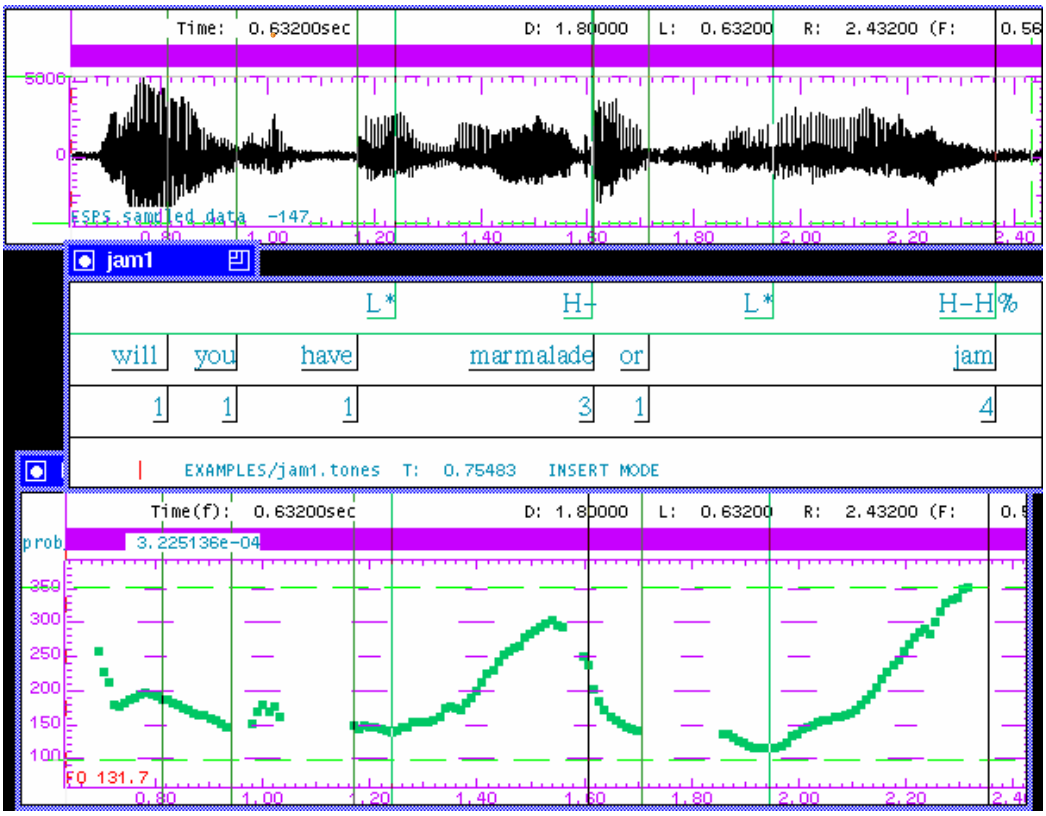


Figure 2.2 An illustration based on ToBI model showing waveforms (upper part), F0 contours (lower part) and the four tiers (in the middle) for tones, words, break indices and miscellaneous events; image adopted from http://www.speech.cs.cmu.edu/tobi/ToBI.1.html

2.2.4 Autosegmental-Metrical theory (AM theory)

AM is a theoretical model that describes the intonation and prosodic structure of languages. The name autosegmental-metrical is derived from two sets of theories. As explained in Gussenhoven (2002) the model is autosegmental because it has a separate tier for segments (vowels and consonants) and another tier for tones (H and L) just like in autosegmental phonology as originally proposed by Goldsmith (1976). It is also metrical because of the assumption that these tiers are hierarchically organized in a set of phonological constituents from segments, syllable, foot, phonological word, phonological phrase, intonational phrase, up to an utterance. According to Ladd (1996, 2008) the theory has its origins from three

influential PhD theses: Liberman (1975), Bruce (1977) and Pierrehumbert (1980). The AM theory sets out to describe both phonological and phonetic accounts of intonation. As Ladd (2008:43) puts it:

> 'the theory adopts the phonological goal of characterizing contours adequately in terms of a string of categorically distinct elements, and the phonetic goal of providing a mapping from phonological elements to continuous acoustic parameters.'

The theory has four basic tenets:

a. Sequential tonal structure

In sequential tonal structure it is assumed that pitch contours are phonologically represented as discrete intonational events. For example, in English, there are two notable intonational events namely, pitch accents and edge tones. To illustrate further on this argument, Ladd gives an example of two echo questions in English as follows (Ladd, 2008:45-46)

A: I hear Sue's taking a course to become a driving instructor.

B: Sue!?

Or

A: I hear Sue's taking a course to become a driving instructor.

B: A driving instructor!?

In both answers the pattern is rise-fall-rise yielding two important intonational events, that is, accentual rise through a prominent syllable (Sue and driv-) as well as a rise on the edge tone. These are the two discrete events as illustrated by the contours in the responses above. In between, there is just a transition between the two events. According to Ladd this is one of the key innovations of AM theory that it draws an explicit distinction between the two events and a mere transition between them. This implies that localized pitch features such as accentual pitch rise, and edge tone rise are linguistically important rather than long stretches of contour.

This tenet works pretty well with intonation languages like English and German which have specified tonal events. In tonal languages, where tones are lexical or grammatical properties of syllables/morae there can be several tonal events than the ones prescribed.

b. Distinction of pitch accent and stress

The second tenet of AM theory is based on the distinction between pitch accent and stress. Pitch accent is defined as a local feature of pitch contour which signals that the associated syllable is prominent. Thus, pitch accents have dual aspects such that they act as building blocks of intonation contours, as well as cueing syllable prominence. It should be noted that not all languages have pitch accents. Moreover, pitch accents have been well established in languages like English and Dutch as intonational pitch accents, as opposed to lexical pitch accents found in languages like Swedish, Slovenian and Norwegian (Ladd, 2008). While pitch accent is viewed as an actual prominence of a syllable, stress is viewed as an abstract lexical property of an individual syllable. This is because there are no clear and/or consistent

phonetic cues for stress. Stress can be realized phonetically in citation forms but unrealized in utterances if a word is not prominent. Therefore, in relation to the sequential tonal structure analyzed in (a) above, the phonological events (i.e. accentual rise on the prominent syllable and edge tone rise on the final syllable) are determined by pitch accents. This implies that pitch accents are important in pitch contours rather than stress.

c. Analysis of pitch levels vs configurations or pitch movements

Another basic tenet of AM theory is to resolve a long-standing controversy between the levels-vs-configurations debate. Ladd points out that the debate has been between those who consider contours as distinctive pitch levels and those who consider contours as pitch movements or configurations. The first thing done in AM theory was to reduce the number of distinctive levels from four (i.e. Low, Mid, High and Overhigh), like that pioneered by Liberman (1975), to two (i.e. High and Low), the number which is also advocated by Pierrehumbert (1980). Pierrehumbert specifically points out that a two-tone theory, which targets levels is attractive because it can describe how the same intonation pattern aligns with different texts by showing the crucial points in a contour and how they align with the crucial points in the text.

It is thus argued that the Hs and Ls are treated as phonological abstractions which can be phonetically realized differently in different contexts when they are scaled in actual F0. Another thing is that the four-level tones can bring about six distinct falling contours, but AM theory reduces all these to one falling contour which is HL, while attributing variations

in the height of a H or a L to other orthogonal factors. Similarly, a 'rise-fall' contour captures all the different configurations.

Furthermore, this tenet advocates an analysis of pitch contours in terms of levels rather than movements or configurations (known as F0 changes in Pierrehumbert 1980). This is said to have been partly influenced by African languages which have falling and rising contours as sequences of level tones rather than units. In such languages, this is clearly observed in tonal processes, particularly when a segment is deleted but its tone is still realized on the following segment. Therefore, in order to capture such tonal orientations, a theory focusing on tonal targets is important than that focusing on tonal configurations. And as Bruce (1977) argues, in the modelling of accentual pitch configurations, reaching a certain pitch target is more important than the movements (rise or fall).

d. Phonological interpretation of global F0 trends

The last tenet of the AM theory is based on the interpretation of global F0 trends. It is argued that overall trends in pitch contours such as gradual lowering of overall pitch range reflect the operation of localized but iterated changes in scaling factors (Ladd, 2008). For example, declination in the AM theory is considered as a gradual lowering in pitch as a result of downstep. Meanwhile, downstep can be viewed as gradually declining trend or a step-wise fall of subsequent highs.

2.3 Intonation in tonal and non-tonal languages

2.3.1 Non-tonal languages

Gussenhoven (2004:12) defines non-tonal or intonation-only languages as the languages without lexical tone. These include languages such as English, German, Greek, and Dutch which have stress instead of lexical tone. One of the earliest works on English intonation is that of Liberman (1975). Liberman points out how different stress placements produce different syntactic structures with different semantic interpretations. He gives an example of the phrase '*English teacher'* which can have stress on the first or second word. When stress is on the first word the phrase means 'a teacher who teaches English', here the compound stress rule (CSR)[1] applies. When stress is on the second word the phrase means 'a teacher who is of English origin', and in this case the nuclear stress rule (NSR)[2] applies. Liberman adds that intonational difference can also be obtained depending on how phrasing is done. This means the location of commas, pauses and intonation breaks play a role in determining the structure and the meaning of a phrase or an utterance. Liberman's approach to English intonation aims at identifying the underlying form of a phrase or sentence, describing its phonetic representations and then determining the interaction between them. His argument is that, in English, text and tune are underlyingly separate. A metrical system then defines the metrical patterns for text and tune independently and then combines them by the establishment of congruence.[3]

---

[1] CSR assumes that the first constituent is strong and thus receives stress (see Booij, 1995).

[2] The NSR re-assigns primary stress to the rightmost primary stressed syllable in a phrasal component (see Chomsky and Halle, 1968).

[3] A set of correspondences between nodes of different metrical patterns.

As already mentioned above, the most remarkable and influential work on intonation is that of Pierrehumbert (1980). Since then, later studies on intonation have followed suit with some modifications, as noted in the ToBI model and AM theory (see below). Pierrehumbert (1980) proposes, specifically for English, that a well-formed utterance is composed of three parts, namely, pitch accent, phrasal accent and boundary tone. The last two together form the edge tones. As stipulated in AM theory, pitch accents and edge tones mark the most important intonational events in non-tonal languages. These tonal events correspond to what Palmer (1922) and O'Connor and Arnold (1973) refer to as head, nucleus and tail.

Generally, the phonetic manifestation of pitch accent is as follows: H* is a relatively high pitch and it may be preceded by a shallow rise; L* is a local pitch minimum low in the speaker's pitch range and may be preceded by a shallow fall; L+H* is preceded by a syllable with a low pitch target followed by a sharp rise to the accented syllable; L*+H begins low within the accented syllable followed by a late rise; H+L* is a step-down from a high to low accent whereby the accented syllable is at the bottom of the speaker's pitch range but preceded by a high pitch target; and lastly H+!H* is a step-down from high to mid accent (Pierrehumbert, 1980; Beckman & Pierrehumbert, 1980, 1986; Grice et al., 2005). Although Pierrehumbert's (1980) version recognizes a seventh pitch accent for English, the H*H, she eliminates it in later modified versions, leaving English with six pitch accents, as in German. Gussenhoven (2005), as well as Arvaniti and Baltazani (2005), on the other hand, suggest five pitch accents for Dutch and Greek.

Furthermore, phrase accents and boundary tones characterize the right edge of an intermediate phrase (ip) and intonation phrase (IP), respectively, in non-tonal languages. While English has two phrase accents, that is H^- and L^-, German and Greek have three phrase accents with an addition of a downstepped or mid high ($!H^-$). Boundary tones are also language-specific and are phonetically manifested at the right edge of an intonational phrase.

2.3.2 Tonal languages

Tonal languages, by contrast, have contrastive tone on every syllable or mora with the purpose of distinguishing lexical items or grammatical functions. This excludes languages like Swedish and Japanese which have tone on specific syllables only. According to Yip (2002) these latter languages are termed 'accentual languages' and are considered a sub-type of tonal languages. Regarding tonal notations there are two main systems used. The first one is the Africanist tradition, which uses the acute accent (á) for high tone, grave accent (à) for low tone and mid accent (ā) for level tone. Low and mid tones can be left unmarked (Yip, 2002:19). Another system is called 'Chao tone letters' based on Chao's work in 1930. This system uses numbers, whereby 1 is the lowest and 5 the highest in terms of pitch. This is particularly used in Asian tonal languages. The Americanist approach uses a similar system to that of Asia but reverses the numbering (Yip, 2002).

Regardless of the differences in the number, location and representation of tones, tonal languages have a lot in common as far as intonation is concerned. One of the compelling issues regarding the intonation of tone languages is how lexical tones interact with intonational tones. As explained by Hyman and Monaka (2011) when the two interact there can be superimposition whereby both tones are visible. Alternatively, there can be avoidance

whereby lexical tones override intonational tones or submission whereby intonational tones override lexical tones. The phonetic correlates of intonation in tonal languages can be grouped into global and local effects. The former includes pitch range expansion (PRE), pitch range compression (PRC), and declination as intonational effects affecting an entire phrase or utterance. Local effects may include phonological downstep, final lowering and boundary effects, as well as segmental effects such as penultimate or final lengthening. It should be noted that these intonational effects can also be found in non-tonal languages although they may be differently implemented. Section 2.5 below is concerned with the phonological representations and phonetic manifestations of different sentence types in tonal and non-tonal languages.

2.4 Intonation and sentence type

2.4.1 Declarative intonation

The first common aspect observed in the intonation of tone languages is downtrend. This is in particular found in tone-terracing languages. According to Clements (1979) tone-terracing, which is widely reported in African languages, displays a regular register shift, and is a process that affects the pitch realization of successive tones. Declarative sentences are reported to have one or more forms of downtrend. These include downstep, declination or final lowering.

To begin with downstep, this is considered a phonological effect from the interaction between H and L tones (Gussenhoven, 2004). Automatic downstep is when there is an

obvious L tone which lowers the subsequent H. Some researchers refer to this as downdrift (Yip 2002, Downing & Rialland, 2016). Non-automatic downstep, on the other hand, is when there is a floating L tone which triggers downstep (Gussenhoven, 2004; Downing & Rialland, 2016). However, it is noted that not all non-automatic cases are caused by a floating L. For example, in Bemba, Kula and Hamman (2016) point out that non-automatic downstep between H tones is due to OCP adherence. Some cases, as argued by Clements (1979), are due to grammatical operations like signalling a syntactic boundary.

Akan (Kwa: Niger-Congo, spoken in Ghana) is a good example of a terraced-level tone language which portrays a downtrend movement following the interaction between L and H tones (Genzel, 2013; Kügler, 2016). Kügler notes that downtrend is also manifested in other types of sentences not only in declaratives. He reports that a simple declarative has downstep shaped by local interaction between L and H and tonal neutralization utterance-finally. Neutralization is manifested when there is a H-tone which becomes L sentence finally, meaning declaratives have L% boundaries even when they end with a lexical H. A combination of downstep and L% boundary for declaratives is also reported in Bantu languages like Bemba (Kula & Hamann, 2016) and Shingazidja (Patin, 2016). Some languages are reported to lack downstep even if H tone is preceded by L tone(s). A good example is Embosi as reported by Rialland and Aborobongui (2016). Although the language has no downstep, a declarative is marked by a boundary L% at the right edge.

In non-tonal languages like German downstep is also observed in established facts alias statements with a pitch contour H+!H*L-L% whereby !H represents a step-down from high

to mid accent followed by a low phrase accent (L-) and a boundary (L%) (Grice et al, 2005). In Greek, however, !H does not represent a downstepped tone but a mid-level pitch (Arvaniti & Baltazani, 2005). Arvaniti and Baltazani argue that the absence of a downstepped H is because phrase accents do not influence boundaries, the interaction which can trigger downstep. English has both downstepped and non-downstepped H tones. The choice of whether to downstep a pitch accent is paradigmatic depending on what the speaker wants to convey. In Japanese (an accentual language) downstep is completely predictable. The predictability is based on the fact that an accentual phrase will be downstepped if the preceding accentual phrase bears an accent and both phrases are in the same intonation phrase (Venditti 1997:129).

Apart from downstep, declination is also reported in some languages. This is a gradual downsloping of the fundamental frequency. It is argued that at the beginning of an utterance the subglottal air pressure will be higher than the end. This is because the speaker slowly eases up on the tension of the muscles he used for the breath intake (Yip, 2002; Gussenhoven, 2004). Declination is reported in both tonal and non-tonal languages although with different mechanisms. In tonal languages, declination is clearly observed in like-tone sequences. In West African languages declination is reported as a phonologized process (Yip, 2002). In Akan, Kügler reports that sentences that contain like tones (all High and all Low) show a similar kind of downtrend as those with an alternating sequence of H tones and L tones. Genzel (2013) as well as Kügler (2016) propose that the observed downtrend in like-tones is a result of a phonologized declination whereby a sentence has an initial high register (h) which is associated with the left edge and a final low register (l) associated with the right

edge. Such kind of presentation has an effect on the global downtrend of F0 over an utterance, as proposed in the AM theory. In Luo, Creider (1978) argues that in declarative sentences the intonation is perceived with a gradual fall in pitch. The fall is clearly heard when the utterance has like-tones such as high tones or low tones. He argues further that in a low tone sequence the fall can be suspended or attenuated when a high tone intervenes. Similarly, in a sequence of falling tones there will be a downstep with a superimposed gradual fall on the overall tone group.

As we have seen above, declaratives are also cued by low endings at the right edge. This trend is called final lowering and it can be found in both intonation and tone languages. Final lowering involves pitch drop to the bottom of the speakers pitch range. According to Pierrehumbert and Beckman (1988) final lowering normally applies in the last-half second or so of a declarative utterance. In African tonal languages, Downing and Rialland (2016) also report that final lowering can range from a single syllable to a phonological phrase. Bemba presents a case where final lowering can affect one syllable or can extend to a number of syllables regardless of whether the final tone(s) is L or H (Kula & Hamman, 2016). There are different ways in which final lowering can be realized in terms of its domain and phonetic implementation (Downing & Rialland, 2016). Gussenhoven (2004) elaborates that final lowering can be grammaticalized whereby some languages have a total downstep of the final H tone (in case a language has non-total downstep in non-final position), some have a falling contour utterance finally, some portray a lack of H-tone spreading utterance finally, others simply downstep the final H regardless of the context on the left. In Akan, as we have already seen, the final H tone neutralizes to a L-tone to mark finality (Kügler, 2016). In some

languages, final lowering is restricted to some tones only. For example, in Mambila, final lowering takes place only when there is a low tone (Connell, 2016). In non-tonal languages like English and German final lowering is also phonetically manifested by L% boundary, which is said to mark completeness or certainty, a typical melody for declaratives. It should be noted that, in non-tonal languages, the L% boundary can be preceded by different pitch contours resulting in other interpretations such as a call or surprise.

Not all languages have L% boundary or final lowering at the end of declarative sentences. Basaa (Makasso et al., 2016) and Kɔnni (Cahill, 2016) do not have final lowering. Meanwhile Chickasaw (Native American) declarative sentences (Gordon, 2005), as well as Greek negative declaratives (Arvaniti & Baltazani, 2005), have boundary H%. Such cases are said to be rare (Gussenhoven, 2004). Additionally, segmental features can also signal the end of a statement or declarative, for example, penult-lengthening in Chichewa and Tumbuka (Downing, 2016), which is also widely attested in Southern Bantu languages.

Downtrends are also reported in the early works of Luo intonation by Tucker and Creider (1975) and Tucker (1994). In their works, they argue that the language has explicit downstep, implicit downstep and grammatical downstep. Explicit downstep corresponds to what has been referred to in this section as automatic downstep, as the preceding L tone lowers the subsequent H. For the implicit and grammatical cases there is no obvious L which lowers the subsequent Hs. This corresponds to what has been presented here as non-automatic downstep. Tucker and Creider elaborate that in grammatical downstep the H tone of the

subject and that of the verb are in a downstep relationship. A detailed discussion about different forms of downstep in Luo is presented in Chapter 3 section 3.3.3.

Furthermore, Creider (1978:329) gives the general characteristics of Luo intonation by providing the phonetic correlates of what he refers to as tone groups. The first one is that a full word stress is found on the last word in a tone group. This is similar to the intonation of languages like English and German which have pitch accents as tonal events at particular points in a phrase or utterance. Second, he points out that there is a slight pause following the stressed item, which is clearly distinguishable from other forms of hesitations or halts. The latter are said to be of short duration with body parts in raised position. Also, while hesitation can be very long, halts are sudden. Lastly, Creider points out that successive tone groups are distinguished from one another by noticeable beginnings and ends.

Nevertheless, Tucker and Creider's analysis of Luo intonation is generic and does not pay particular attention to specific syntactic structures. Furthermore, nothing has been said regarding intonation at the final positions of phrases or utterances. This creates a gap that this thesis aims to fill. An in-depth investigation of Luo declaratives is conducted and the following questions are addressed: How are declaratives cued in Luo? Are there any forms of register shifts as observed in other tonal languages? Does Luo have final lowering? If yes, how is it grammaticalized? Do all four tones exhibit a final lowering effect or is it restricted to some tones?

2.4.2 Question intonation

In this section, a survey of question intonation is explored. As it was for declaratives, there are language-specific differences in how questions are intonationally cued. While a falling

pitch is generally associated with statements, a rising pitch is associated with questions (Ohala, 1983). But this is not always the case as we will see below. To begin with content questions, some languages show a similar intonation contour to that of statements. This means they also have a boundary L% at the right edge. This is reported in non-tonal and tonal languages alike. But some languages like Greek have a boundary H% for wh-questions (see Arvaniti & Baltazani, 2005) while others, like Fipa (see Riedel & Patin, 2011) have a boundary H(L)%. High scaling of questions is also reported particularly in tonal languages. For example, wh-questions in Bemba and Chichewa are produced at a higher pitch register compared to statements.

Polar questions, also known as yes-no questions, are the most diverse in how they are phonetically manifested. In most intonation languages, polar questions have a final rise which is attributed to a boundary H% at the right edge. For example, English has the contour pattern L* H- H% with a rise at the end (Pierrehumbert, 1980). Similarly, in German, neutral yes-no questions have a rising contour represented by L*H-^H% with an up-stepped high tone marked by (^) (Grice et al., 2005). Contrary to other intonation languages, Chickasaw has a boundary L% tone for yes-no questions (Gordon, 2005). Gussenhoven (2004:91) adds that, apart from high peaks or high ends, late peaks can also indicate questions. He gives an example of Hungarian where earlier peaks are attributed to declaratives but later peaks to interrogatives.

In tonal languages H% boundary is reported in Mandarin and Cantonese. African tonal languages portray a wider variation in the way polar questions are produced. In her survey of

polar question prosody in African languages, Rialland (2007: 35) reports that there are languages which exhibit high pitched intonation while others exhibit non-high pitched intonation which she terms ‘lax prosody’. The former group, like most intonation languages, have a rising intonation as a cue for polar questions. Other strategies in this case are cancellation or reduction of downdrift, cancellation or reduction of final lowering and HL melody (Rialland, 2007). For the case of lax prosody, a falling intonation, mid or polar tone, lengthening, breathy termination and open vowels can signal yes-no questions. Falling intonation, for example, is reported in Akan and Bemba. In Akan, Kügler (2016) reports that yes-no questions together with L% boundary have intensity increase and lengthening phrase finally. In Bemba, the L% boundary in yes-no questions targets the final syllable rather than a string of syllables as in statements (Kula & Hamann, 2016). In this language, another salient intonational feature which distinguishes statements from questions is pitch range expansion (PRE) whereby the lexical tones are produced at a higher register. Chichewa (Downing, 2016) and Shingazidja (Patin, 2016), on the other hand, have HL% boundary over the final two syllables or the last syllable respectively, creating a rise-fall F0 movement. Additionally, Chichewa and Shingazidja suspend downstep, a tendency also observed in Bemba. As reported earlier, segmental strategies can also cue yes-no questions. This is exemplified by vowel lengthening in Akan and Konni and suppression of penultimate lengthening in Tswana (Downing & Rialland, 2016).

The phonetic correlates of echo questions are also language-specific but generally manifested within the lines of content or yes-no questions. In Mandarin and Chickasaw echo questions have overall pitch register raising, with Chickasaw portraying a H% boundary. An English

echo question is said to have a rise-fall-rise contour yielding two intonational events at the prominent syllable and at the edge tone (Ladd, 2008). Palmer (1922) identified a high rising for both echo questions and echo statements. In German, Grice et al. (2005) report a rising contour for echo questions represented as L*H-^H% whereby the final high tone is upstepped, an intonational contour similar to that of wh-questions. In Shingazidja echo questions have LH*!H% contour where the final tone is downstepped (Patin, 2016).

The literature on question intonation reveals different phonetic cues for tonal and non-tonal languages. For the case of interrogatives in Luo, Creider (1978) points out that the intonation is similar to that of a declarative but produced at a higher pitch range. However, interrogatives have greater pitch fall. Rialland (2007) places Luo under high pitched prosody because it has register expansion and downdrift/downstep reduction. This information alone does not suffice to describe Luo question intonation, for both content and yes-no questions. Again, nothing is said on what happens at the boundaries of questions, which also makes the topic worth investigating. Therefore, another question which will be addressed in chapter five is what are the intonational cues for content and yes-no questions in Luo?

2.4.3 Focus

According to Güldemann et al. (2015) information structure is about how speakers structurally encode propositional content with respect to their assessment of knowledge that is shared (or not) by interlocutors in a particular communicative situation. Krifka (2008) points out further that information structure includes givenness, focus and topic whereby new information is the complement of given, background information the complement of focus and comment the complement of topic. In this section, attention will be given to focus and

topic. There are different ways in which focus and topics are presented cross-linguistically. This ranges from morphological marking where markers are attached to the topicalized or focused constituent to syntactic manipulation of a sentence such as dislocations and clefts (Güldemann et al. 2015). Some languages have prosodic marking such as change in pitch or duration. Downing and Rialland (2016) add that focus can also be marked through demarcation. This involves placing a juncture before or after a focused constituent. There are broadly two main types of focus. The first one is an answer to a wh-question[4]. Inherently, wh-words are said to be focused. This means logically the wh-word itself is the focus of the question (Ladd, 1996). This type of focus is called presentational focus (Selkirk, 2002; Gussenhoven, 2004) or non-contrastive focus (Hartmann, 2008). The second type is contrastive focus. This includes alternative focus, and correction focus, among others. It can be obtained through exhaustive questions.

In non-tonal languages, it is generally assumed that focused constituents, alias new information, is prosodically more prominent than the background information. This means focused information is produced with a higher F0 value (Liberman and Pierrehumbert, 1984). According to Gussenhoven (2002, 2004) focus involves the grammaticalization of the effort code. This is because wider energy excursions are applied when a constituent is focused. Although focus prominence is believed to be a common intonational tendency at least in non-tonal languages, Ladd (1996:171) elaborates that English does not accentuate question words unless they occur as echo-questions in a non-wh-movement as in '*You did WHAT*?'. Such prominence, can be found in Romanian which accentuates the wh-words as in *UNDE mergi*?

[4] Here 'wh-question' is used as a cover term for content questions.

‘WHERE are you going?’ yielding a H* L L% contour (Ladd, 1996:173). Moreover, in most non-tonal languages an answer which corresponds to the wh-word is phonetically accentuated. This is represented by L+H* pitch accent for narrow focus or a monotonal H* for broad focus. The former is when only a part of a sentence is in focus while the latter is when the whole utterance is in focus.

In tonal languages, such as Mandarin, the prosodic marking for focus is somehow similar to that of non-tonal languages. Narrow focus in Mandarin has an expanded pitch register while the non-focused constituent exhibits pitch compression. Broad focus, on the other hand, does not have pitch register expansion nor compression (Peng et al., 2005). In the recent studies of intonation in African languages, Downing and Rialland (2016) point out that several African languages do not have focus prosody. Only Akan (Kügler, 2016) and Shingazidja (Patin, 2016), among the twelve languages in the volume are reported to mark focus. These languages mark focus by demarcative means not by stressing or making it prominent. Such languages use prosodic phrasing to mark the presence of a focused (or topicalized) constituent (Güldemann et al., 2015; Costa & Kula, 2008; Patin, 2008; Downing et al., 2004). This means a boundary is inserted at the edge of a focused constituent. This is clearly illustrated below from Shingazidja (Patin, 2016: 3-4).

19. [(ŋgam-andzo tʃáy)*Φ*]IP

1.PER-like 9.tea

‘I like tea’

20. [(ŋgam-andzó)*Φ* (tʃa!í)*Φ*]IP

‘I LIKE tea

Focused information in (20) shows a boundary at the right edge of the VP marked by (*Φ*). This phrasing is different from that of a non-focused VP in (19). This shows that (19) has one prosodic phrase (VP + OBJ) and (20) has two prosodic phrases (VP) (OBJ). The demarcation in (20) is cued by lack of H-tone shift to the following word. This is contrary to what is observed in (19), where the H-tone on the VP shifts right-wards to the following object.

Previous studies on focus in Luo are based on syntactic orientation and they mostly feature cleft constructions. This is because subjects, as in other languages, cannot be focused sentence initially unless they are clefted. Okoth-Okombo (1997) is one of the researchers who ventured into focus marking in Luo. Okoth-Okombo argues that focus in Luo is marked by a particle *-e*. In a later study by Cable (2012) another particle termed the normal copula *en-* was added. The general rule is that the normal copula *en-* precedes the XP while the focus copula *-e* follows the XP, and not vice versa. Normal copula *en-* is also used in fronted question words. (A detailed discussion of cleft constructions is given in chapter 6 section 6.3). A contribution to this discussion will be made by investigating the phonetic manifestation not only of cleft constructions, but also of other forms of focus such as contrastive and non-contrastive focus. The questions that will be addressed include: Does Luo prosodically mark focus? If yes, how is it cued?

2.5 Prosodic structure

This section focuses on prosodic constituents and their relation to syntactic constituents. It starts by giving the background information on prosodic constituents and moves on to elaborate on language-specific correlates of prosodic structures, focusing on the higher levels of phrasing like phonological phrases and intonation phrases.

In the previous sections, we have seen the role of prosody in distinguishing sentences and other grammatical categories. Another function of prosody is to divide up a stream of speech into chunks or breathe groups (Ladd, 2008). This division leads to the use of the term prosodic phrasing. It is assumed that a prosodic structure consists of categories which are hierarchically arranged into layers. According to Selkirk (1986) and Nespor and Vogel (1986) the basic prosodic units are a syllable (σ), foot (φ), phonological word (Pw, ω), phonological phrase (PhP, *Φ*), intonation phrase (IP) and an utterance (Utt, U). Selkirk (1986) proposes the syllable as the lowest prosodic category which is dominated by the foot which in turn forms prosodic words. Several prosodic words form a phonological phrase. An intonation phrase can comprise one or more phonological phrases. Lastly, an utterance can constitute one or more intonation phrases. On the other hand, Hayes (1989) proposes five levels. These are the utterance, the intonation phrase, the phonological phrase, the clitic group and the word. The clitic group is also present in Nespor and Vogel's (1986) prosodic hierarchy. In Hayes (1989) prosodic hierachy the word is considered the lowest category in a hierarchy of prosodic structure, where numerous phonological rules such as stress assignment or vowel harmony are bounded. The clitic group comprises of a single content word together with all contiguous grammatical words in the same syntactic constituent. Thus, every content

word belongs to a separate clitic group (Hayes, 1989: 207-208). Furthermore, one or more clitic groups can form phonological phrases. The phonological phrase in turn forms an intonation phrase (IP). Earlier studies of intonation, particularly in English, considered the IP as the only constituent (See Pierrehumbert 1980). Later on Pierrehumbert and Beckman (1986, 1988) proposed another level known as intermediate intonational phrase (ip), which comprises pitch accent and phrasal accent but no boundary tone. In other languages, particularly African languages, an ip corresponds to a phonological phrase or minor phrase (Selkirk, 1996, 2002). The term intermediate phrase (ip) is widely used in intonation analyses which follow the ToBI model and correspond to a level 3 Break Index. Pierrehumbert and Beckman (1988) in their analysis of Japanese consider this lower level of phrasing as an accentual phrase (AP).

Numerous studies have been conducted to determine the levels of phrasing in different languages. In intonation studies interest has been on the higher levels of phrasing that is above the phonological word such as phonological phrases (alias ip, AP) and intonation phrases (IP). There is a greater diversity of views regarding the representation of these higher prosodic structures (Truckebrodt, 2007). Some authors (Odden, 1987; Truckenbrodt, 1995, 2007) argue that prosodic structures are directly related to syntactic structures. Selkirk (2000), on the other hand, argues that prosodic phrasing is independent of, but related to syntactic and information structure of a sentence. Truckenbrodt (1999, 2007) argues that phonological phrases relate to syntactic XPs such as Noun phrases (NPs), Verb phrases (VPs) and Adjective Phrases (AdjP) while intonation phrases relate to syntactic clauses.

Furthermore, edge-alignment theory (see Selkirk, 1986, 2002; Truckenbrodt, 2007) stipulates that the edge of a syntactic structure should align with the edge of a prosodic structure. This means the right and/or left edge of an XP should be aligned with the phonologically cued prosodic domains. Truckenbrodt (2007) proposes a constraint Align-XP,L/R that forces syntactic phrases to conform with phonological phrases. Truckenbrodt (2007), however, argues that the forming of phonological phrases can be blocked by a constraint Wrap-XP which requires every XP to form its own phonological phrase. This prevents splitting of XPs into multiple phrases. Similarly, in Match theory (Selkirk 2011), a syntactic clause is matched by an intonation phrase, a syntactic phrase by a phonological phrase and a syntactic word by a phonological word. This means, in Match theory, phonological domains mirror syntactic constituents, suggesting that Match theory is more restrictive than Align/Wrap theory (Selkirk, 2011).

There is also a great diversity in how XP edges are phonetically cued cross-linguistically. Most, non-tonal languages such as English, German, Dutch and Greek have two levels of phrasing. German, for example, has the intermediate phrase (ip) and the intonation phrase (IP) (Grice et al., 2005). An ip, as already explained above, contains one pitch accent followed by a contour from the last pitch accent to the end of a phrase. German has three ip edge tones namely L-, H- and !H- which align with the post-nuclear stressed syllable. An IP, which is of the highest level in the prosodic hierarchy, combines the edge tone from the preceding ip and a boundary H% or L%. Different combinations yield different pragmatic interpretations. On the other hand, Dutch prosodic phrases are cued at both left and right edges (Gussenhoven, 2005). This means the language has an initial boundary tone (%T) and

a final boundary tone (T%). Three initial boundary tones at the beginning of an IP are low or mid pitch, high pitch and falling pitch represented as %L, %H and %HL, respectively. Meanwhile, there are only two final boundary tones, high (H%) or low (L%) at the right edge of an IP. Gussenhoven (2005) argues that the final tones are optional signaling an absence of tone at the IP right edge. Gussenhoven notes that Dutch has no intermediate phrases as in German, but rather the phonological phrases (PhP) can directly form an intonational phrase (IP). Note that phonological phrases are mostly reported in African tonal languages.

Other languages have accentual phrases (APs) instead of phonological phrases. Gordon (2005) proposes three levels of phrasing in Chikasaw. These are the intonation phrase (IP), the accentual phrase (AP), and the prosodic word. An IP contains one or more APs. Gordon argues that it is important to have the prosodic word level in order to distinguish the domain for word-level stress assignment and the rhythmic lengthening domain which is sensitive to morphology. The rationale for this is that Chickasaw has two types of pitch accents, the morpho-lexical pitch accent and the nuclear pitch accent. The former falls on a single syllable of some verbs while the latter is highly predictable as it falls on the final syllable or heavy syllable of the final word in an intonation phrase. In Japanese, Pierrehumbert and Beckman (1988) propose three levels of phrasing, the accentual phrase (AP), the intermediate phrase (ip) and the intonation phrase (IP). Venditi, on the other hand, argues for two levels in Japanese excluding the ip in the Pierrehumbert and Beckman (1988) analysis. She argues that Pierrehumbert and Beckman's experiments focused on read-text phrases lacking phrases from spontaneous speech. Japanese APs are cued at the left and right edges by delimitative phrase tones. These are initial rise at the left edge and a fall at the right edge, annotated as

%L H- and L%, respectively. An example of Japanese phrasing is given below, where an IP may contain one or more APs.

21. [{*sa'Nkaku no*}$_{\mathbf{AP}}$ { *ya'ne no*}$_{\mathbf{AP}}$]$_{\mathbf{IP}}$ [{*maNnaka ni*}$_{\mathbf{AP}}$]$_{\mathbf{IP}}$ [{*okima'su*}$_{\mathbf{AP}}$]$_{\mathbf{IP}}$

triangle-GEN roof-GEN middle-LOC put

'I will place it right in the centre of the triangle roof.'

(Slightly modified from Venditti, 2005:176)

Additionally, Venditti proposes three types of boundary tones, which she calls boundary pitch movements (BPM). These tones are rising H%, rise-fall HL% and a scooped LH% found at the right edge of an IP. The H% boundary, which in the previous analyses indicates questions, can be divided into two depending on their F0 heights. There is a high-rise cueing question interpretation and mid-rise cueing insistence. The HL% boundary is said to be found in casual speech of young speakers and it sounds 'explanatory'. The scooped LH% gives an incredulity interpretation. At the left edge, an IP is signaled by pitch reset. Like Japanese, Korean also has APs and IPs. While an AP is demarcated both at the left and right edge, an IP is cued utterance finally with at least nine boundary tones (Jun 2005). Jun elaborates further that the left edge of an AP in Korean can either be H or L depending on the laryngeal features of the initial segment. Meanwhile, the right edge has fourteen possible patterns (Jun 2005: 214).

Turning to tonal-languages, most of them portray two levels of phrasing, the phonological phrase and the intonation phrase. However, some languages are reported to have one level of phrasing. If a language happens to have one level of phrasing then it will be an IP. An example of one-level phrasing is from Chichewa as argued by Downing (2016). In the previous study of Chichewa by Kanerva (1990) the language is reported to have two levels of phrasing, PhP and IP. Karneva argues that the PhP is determined by penult lengthening phrase finally and other tonal processes such as high tone spreading. On the other hand, Downing (2016) refutes the two levels and suggests a one-level prosodic phrasing, the IP. According to Downing an IP in Chichewa aligns with syntactic clauses since the intonational tones seem to target the right edge of clauses and initial topics. The VP and all its complements can form a single phrase.[5] This is unlike Tumbuka (a prosodically-related language, also spoken in Malawi) where only the first complement is phrased with the verb. Downing also argues that in Chichewa subjects are phrased separately only when they are topicalized. Thus, a Chichewa VP like the one presented in (22) would be bigger than expected for just a lower level phonological phrase.

22. (a-na-mény-á nyumbá ndí mwáála)
    1SBJ-RECENT.PST-hit 9.house with 3.rock
    'S/he hit the house with a rock.'

(Kanerva 1990:98)

[5] One explanation offered for this disparity is that Kanerva (1990) and Downing (2016) investigate different dialects of Chichewa.

For Tumbuka, Downing (2016) argues that a phonological phrase is determined by vowel lengthening and a H-tone on the first mora of the lengthened vowel. Subjects and topics form distinct phonological phrases while the verb and its first complement form another phrase followed by other complements which also form a distinct phrase, as illustrated in (6). Both PhPs and IPs are cued by penultimate lengthening although the lengthening is longer in IPs than in PhPs.

23. (Subj) (V NP) (PP)

(m-nyamáata)	(wa-ka-timba	nyúumba)	(na	líibwe)
1-boy	1SBJ-TAM-hit	9.house	with	5.rock

'The boy hit the house with a rock.'

(Downing, 2016:4)

Bemba is also a tonal language with two levels of phrasing, with both PhPs and IPs. According to Kula and Bickmore (2015) and Kula and Hamann (2016), in Bemba, a PhP is cued by rightward High Tone Spreading (HTS). When HTS occurs unboundedly, it indicates a phonological phrase boundary but when it occurs boundedly it implies lack of such a boundary. An intonation phrase, on the other hand, is cued at the left edge by pitch reset and at the right edge by a L% or H% boundary tone (Kula & Hamann, 2016). The following examples portray possible phonological phrasing in Bemba.

24. (bá-ká-mú-lóóndólól-á)PhP (Bùùpé)PhP
2SM-FUT3-1OM-explain-FV Bupe
'They will introduce him, Bupe.'

25. (bá-ká-lóòndòlòl-à Bùùpé)PhP
2SM-FUT3-explain-FV Bupe
'They will introduce Bupe.'

26. (bá-ká-mú-shíík-íl-á)PhP (Chítúúndú)PhP (cáángá)PhP (bwíínó)PhP
2SM-FUT3-1OM-bury-APPL-FV Chitundu 1a.bushbaby well
'They will bury the bushbaby for Chitundu well.'

27. (bá-ká-shíik-ìl-à Chìtùùndù)PhP (càángà)PhP (bwìinò)PhP
2SM-FUT3-bury-APPL-FV Chitundu 1a.bushbaby well
'They will bury the bushbaby for Chitundu well.'

(Kula & Hamann, 2016:5)

Kula and Hamann (2016) point out that the difference between (24) and (25) is that (24) has an object marked verb which prompts the complement to be phrased independently from the verb. Meanwhile (25) has no object marker in the verb hence the complement is phrased together with the verb. Furthermore, H-tone spreads unboundedly in (26) affecting all the words in a sentence. Note that this happens when the object marker is present in the verb. On

the other hand, H tone spreading stops in (27) where lack of an object marker is seen as prompting the complement to be phrased together with the verb.

In order to examine the levels of phrasing and their respective phonetic cues in Luo, sentences with different constituents (i.e. short versus long, simple versus complex) will be investigated. The analysis will be presented in chapter 7.

Chapter Three

Tonal Structure

3.0 Introduction

The aim of this chapter is to analyse the tonal structure of Luo based on the current methodological approaches, in this case the F0 scaling. The analysis will include an examination of lexical/grammatical vs derived tones, as well as underlying vs. surface tones. In doing so we will see the extent to which previous studies on Luo tonal structure by Creider (1978), Omondi (1982) and Tucker (1994), which were entirely based on the auditory and musical approaches, conform to the current methods and theories.

Generally, like in most tonal languages, tone in Luo has both lexical and grammatical functions. This means a word can change its lexical meaning or grammatical function by changing its tonal specification. This is illustrated in (28). The disyllabic word *pala* in (28a) has lexical L tones to mean 'knife' but when there are H tones the word changes its meaning to 'ochre' as in (28b). In (28c) the word *otedo* with HHL tonal patterns shows progressive aspect while that in (28d) with L HL L tonal pattern shows completive aspect.

28. (a) pala 'knife'
    (b) pálá 'ochre'
    (c) ótédo 'He is cooking'
    (d) otêdo 'He has cooked'

In section 3.1, I argue that Luo has four lexical tones, that is, two level tones (H and L) and two contour tones (HL and LH). Apart from lexical tones, the section also provides an analysis of grammatical tones, which are of vital importance in the structure of the language. Section 3.2 provides the possible tonal patterns in the language. In section 3.3, I elaborate on the contexts which produce derived tones through some tonal processes. The chapter ends by analysing the possible tonal typology for Luo based on the criteria proposed by Hyman (2001, 2010).

3.1 Tone inventory

In the literature on Luo, there is a huge disparity on the exact number of tones available in the language. Different scholars argue for a varied number of tones from two to four to several tones. The central works on Luo tone, as already mentioned above, include Creider (1978), Omondi (1982) and Tucker (1994). Creider (1978) identifies three basic tones, namely High, Low and Falling. He also has downstepped tone as derived tone which occurs after a high or a falling tone. Omondi (1982) argues for two basic tones, High and Low and two derived tones, Rising and Falling. Tucker (1994), on the other hand, argues for several tones (i.e. High level, Extra high, Downstepped high (Mid), Low, Low descending, High falling, Long high falling, Extra high falling, Rising, Falling-rising, Rising-falling and compound tones). His analysis of tones includes both lexical and post-lexical tones, as we will see later in this discussion.

These disagreements on the number of tones might be due to different methodological approaches used by the researchers. While Omondi's approach is auditory, Tucker and Creider used the so called tonometer [6] and pitchmeter which are in essence musical. Moreover, Tucker (1994) admits that the solfa renderings from the tonometer readings reflect tonemes rather than tones and are not tied to absolute pitch levels (p. 344). Currently, there are more digitized methods for obtaining pitch tracks which are more convenient than musical or auditory means. Thus, the findings on Luo tones presented in this chapter were obtained by investigating their respective acoustic signals from digital pitch tracking. In this form of measurement, when the frequency of the vibration of the vocal cords is higher, it yields higher fundamental frequency (also known as F0) of the acoustic signal (Gussenhoven, 2004). Usually, F0 is expressed in hertz (Hz), as the number of periods per second (ibid). Thereafter, the pitch tracks were plotted against a time scale to provide a visual representation of the F0. The visual representations were then used to argue for the presence of a particular tone. Most of the words used were replicated from previous studies. This was done purposely to see the extent to which the current findings differ from or conform to the previous ones.

Before investigating the types of tone in Luo, let me explain the location of tone in the language. In autosegmental theory by Goldsmith (1976) it is proposed that every tone must be associated to some tone-bearing unit (henceforth TBU) and every TBU must have a tone, at least on the surface. The TBU can be a syllable or mora. Yip (2002) also notes some cases

[6] This is a musical instrument which consists of a comb containing 56 lamellae against a background screen giving the musical value of each lamella. The speakers speak through a microphone and the appropriate lamellae buzz, their place on the scale being noted at the time (Tucker 1994:323). The pitch values from the tonometer are labelled in numbers below the sentence.

where tones can be directly associated to the segments. For the case of Luo, a syllable is the TBU. According to Yip (2002) one of the diagnostics for a syllable being the TBU is that the language should allow contour tones on both light and heavy syllables. This means two autosegmental tones are associated to a syllable and not the mora, the case which is also observed in Luo.

3.1.1 Low vs High tone

The first type of tone reported in Luo is a low tone (L-tone). This tone can be found in monosyllabic, disyllabic or multisyllabic words. Tucker (1994) provides numerous words that he considers to be low-toned. Here I have chosen the monosyllabic word *pi* 'water' and the disyllabic word *mbaka* 'conversation/story', to use as examples. The acoustic signals for these words are shown in Figure 3.1 and 3.2, respectively.

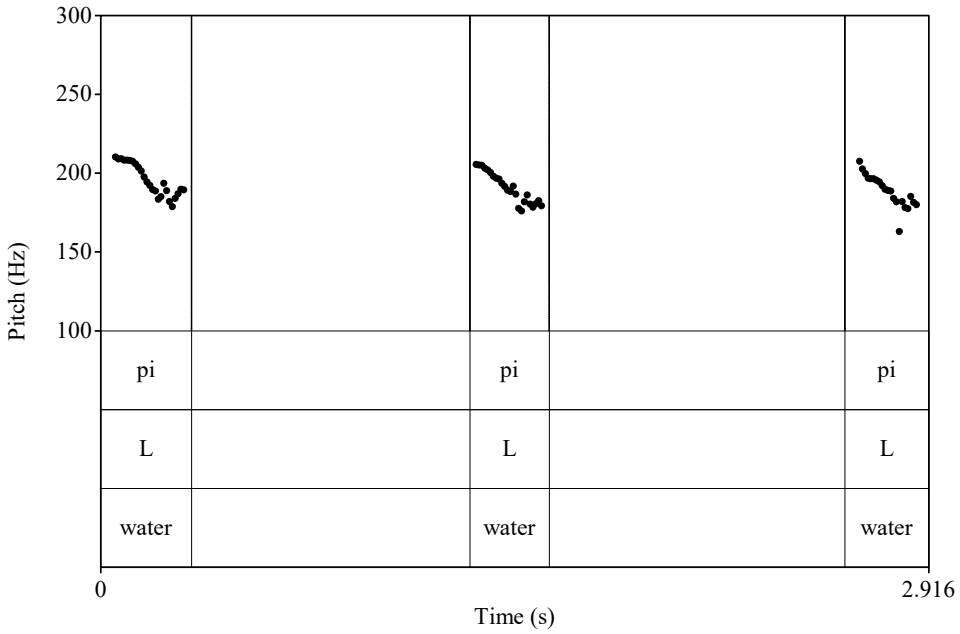


Figure 3.1 Three repetitions of an L-toned word with one syllable

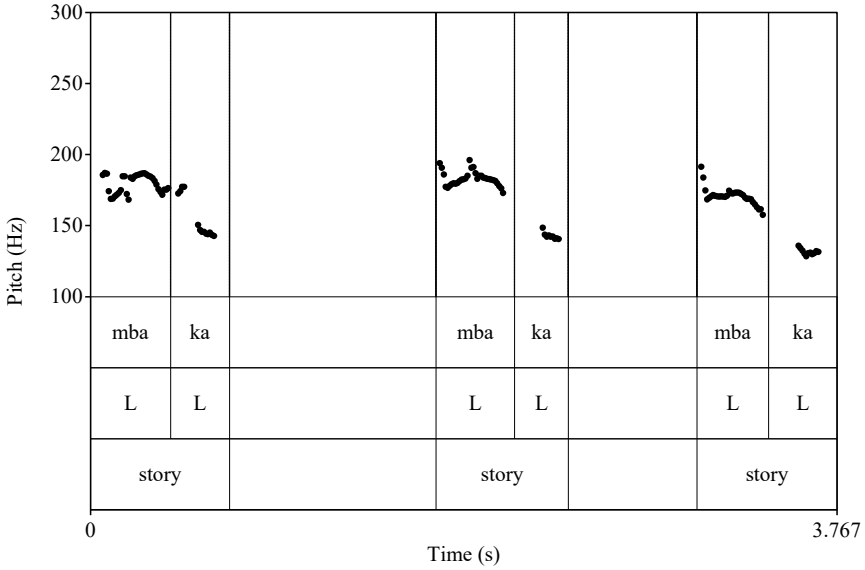


Figure 3.2 Three repetitions of a disyllabic word with L L tonal pattern

First, it is observed that low-toned words, when in isolation, have a falling effect indicating a drop of pitch as illustrated in Figures 3.1 and 3.2. In these figures, it is apparent that this falling effect makes them appear as falling contours rather than level low. Yip (2002) argues that it is normally confusing to decide whether a particular tone is truly a level or a contour tone if it has such effect. However, she argues that in case the fall is small and phonologically irrelevant, that particular tone should be treated as a level rather than a contour tone. This is because when producing a very low-toned word it takes time for the voice to drop to the lowest pitch (i.e. production effect). In order to test the phonological (in)significance of the fall in low tones the differences in pitch values between a word with low tone and a word with falling tone were compared. The results show that the fall of a falling contour is approximately twice the fall of a L-tone. This will be further explained in section 3.1.2 under contour tones. Tucker (1994) in his analysis refers to the fall in low tones as a downdrift effect. In this study, this phenomenon has been observed in all L-toned words in isolation

regardless of the number or structure of the syllables. This means both monosyllabic and multisyllabic words, as well as open and closed syllables, can have the falling effect. Tucker (1994) adds that the falling cadence depends on the speaker, occasion and context. For example, male speakers may portray a steeper fall than female speakers.

Moreover, the language has level low observed in a sequence L H as in Figure 3.7. In this figure, the disyllabic word *sabûn* 'soap' has level low tone on the first syllable followed by a HL contour. Likewise, when a low-toned word is put into larger units such as phrases or sentences it is also portrayed as a level tone. Such examples are found in the subsequent chapters where different types of sentences are discussed.

Apart from low tone, Luo is also said to contrast H-tone. A monosyllabic word *món* 'women' and a disyllabic word *náŋgá* 'cloth' have been used as examples and are illustrated in Figure 3.3 and 3.4, respectively. While low-toned words, as portrayed above, have a falling effect, H-toned words in these figures are phonetically manifested with a sustained level syllable-finally. This means the pitch does not fall to the bottom of the speakers' pitch range as seen in the low-toned words *pi* and *mbaka* above.

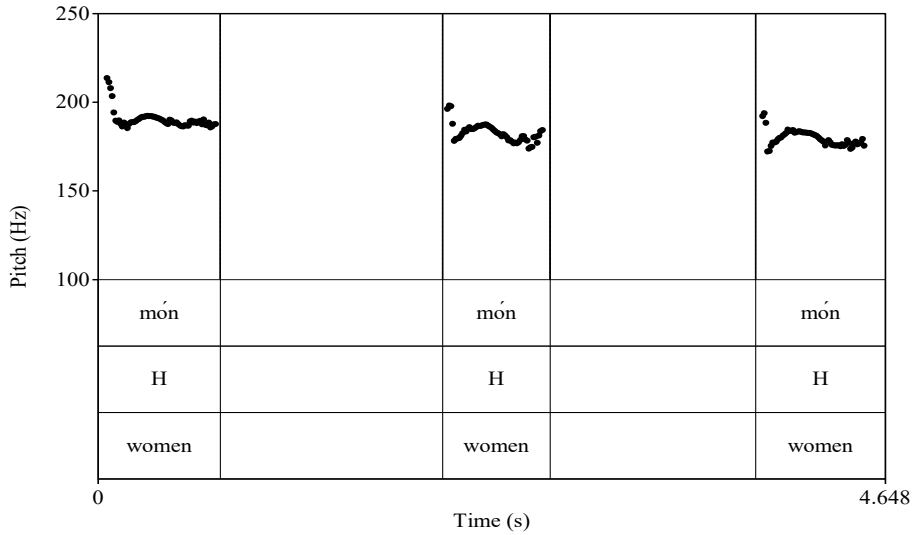


Figure 3.3 Three repetitions of a H-toned word with one syllable

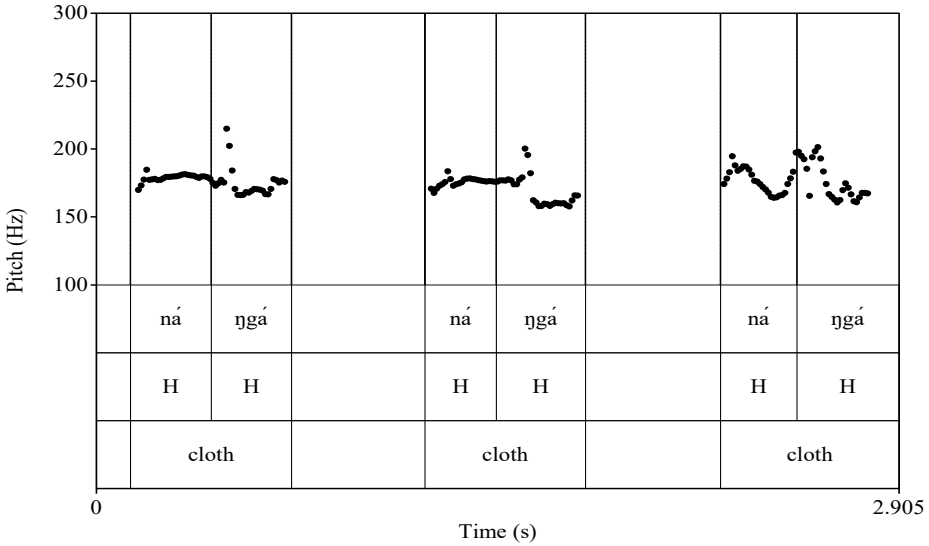


Figure 3.4 Three repetitions of a disyllabic word with H H tonal pattern

The distance between a L-tone and H-tone is Luo is subtle. The F0 contours observed in Figures 3.1 to 3.4 show that both H-toned and L-toned words are produced between 150Hz and 200Hz. Maddieson (1978) argues that a language which contrasts two level tones has a

smaller distance between the two, compared to a language which contrasts three or four level tones. However, there is still a notable difference observed between the two types. The averaged F0 values computed from four speakers from 12 tokens (i.e. three repetitions per participant) show that H-tones are relatively higher than L-tones. This is clearly illustrated in Figure 3.5. First, the difference in pitch between the initial syllables of the L-toned word *mbaka* and the H-toned word *náŋgá* is around 10Hz while that of the second syllables is 36.8Hz. For the reasons already explained above, the latter case shows a greater falling effect. Secondly, on average, the L-tone is produced between 154.89Hz and 118.97Hz while H-tone is produced between 164.87Hz and 155.8Hz.

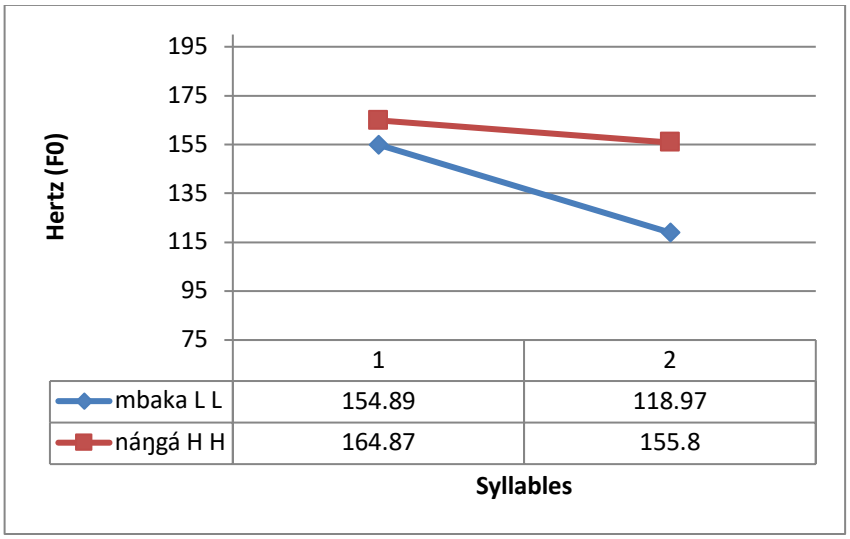


Figure 3.5 Average F0 values of H-toned word (red) vs L-toned word (blue) aggregated from four speakers

Besides L and H, Tucker (1994) also presents another type of H-tone, namely extra high tone (marked by double acute accent ő). According to Tucker, extra high tone occurs in non-final position when the tonal pattern HLH is followed by another H or a downstepped H. He provides examples reproduced in (29), amongst others. In these examples the so-called extra high tone only appears in phrasal context aiming at producing different pragmatic meanings

as in (29a) and (29b). In this case the extra high tone seems to be post-lexical rather than lexical and hence does not qualify as an additional lexical tone.

29. (a) gí-ðǐ́ kod-é
SM3PL-go with-OM3SG
'Let them go with it.'

(b) gí-ði kod-é
SM3PL-go with-OM3SG
'They have gone with it.'

(Tucker, 1994:325)

It is also a matter of debate whether Luo has another type of lexical tone known as a downstepped H. Tucker refers to this tone as a built-in downstep and sometimes calls it mid-tone. According to Tucker, built-in downstep can be lexical or syntactic. The former is the point of concern in this section. Tucker provides examples reproduced in (30) below where H!H[7] tonal pattern in (30b) yields a different but related meaning from the HH tonal pattern in (2a). Using tonometer readings he explains that the solfa rendering for (30a) is r:r while that of (30b) is f:d. This indicates that the tone of the second syllable in (30b) is lower than that of the first syllable. He, however, concludes that these solfa renderings do not reflect absolute pitch levels as they reflect tonemes and not tones. Additionally, Creider (1978) as well as Omondi (1982) argue that downstepped H is predictable as it appears after a high or Falling contour. This makes the downstepped H a derived tone rather than a lexical tone.

[7] It should be noted that Tucker (1994) uses an apostrophe on top of a vowel to mark downstepped H instead of an exclamation mark (!). Thus the word *sí!ró* is annotated as *sírò* in Tucker's analysis. An exclamation mark is used in current autosegmental theories.

30. (a) síró ‘a prop’

(b) sí!ró ‘propping’

(Tucker, 1994:344)

Given the examples above, there is no strong evidence that extra H tone and downstepped H occur as underlying lexical tones. I will therefore consider H and L to be the only contrastive level tones in Luo, whereby a L-tone can be manifested with a falling effect. An effect which is phonetic. Following the Autosgmental-Metrical theory, which is described in chapter 2 section 2.3.3, H and L tones will be treated as the only distinctive level tones which are realized differently when scaled on the actual F0.

3.1.2 Contour tones

In African languages contour tones are treated as sequences of level tones rather than units (Yip, 2002; Ladd, 2008). This is contrary to most Asian languages such as Mandarin and Cantonese which treat contour tones as unitary elements, which cannot be decomposed into level tones. Thus, in African languages H and L can co-occur on the same syllable and yield a phonetically falling or rising contour. Ladd (2008) points out further that the motivation for analysing contours as sequences of level tones rather than unitary elements is based on segmental elision. This means that when a segment, usually a vowel, is elided its tone can still be realized and carried over by a preceding or following syllable.

Assuming that the contours are made up of level tones and the number of levels is two, then the expected contours are HL fall and LH rise (Yip, 2002). Similarly, since Luo has two levels two types of contours are expected. These are falling (HL) and rising (LH) contours. In

the previous literature these contour tones have been portrayed differently by different authors. Creider (1978) argues that the falling contour is one of the basic tones in addition to H and L tones. Creider does not report the rising contour either as a basic or derived tone. Omondi (1982), on the other hand, argues that rising and falling contours are derived tones obtained after tonal contraction or tonal assimilation. These tonal processes will be described later. Despite these arguments, Omondi (1982), however, still provides examples where the falling contour occurs as an underlying lexical tone in a non-derived context. Contrary to Omondi (1982), Tucker (1994) gives extensive examples where both falling and rising contours occur as lexical, derived as well as post-lexical tones.

Yip (2002) argues that contour tones can be simple or complex. Simple contours involve a one-way fall or rise while complex contours start with a rise followed by a fall or a fall and then a rise creating convex and concave shapes, respectively. It should be noted, however, that the simplicity or complexity of a contour shape might be affected by other segmental factors such as voicing and/or sonority. Thus, the distinction between simple and complex contours is simply based on the observed F0 contour. To begin with falling contours, let us consider the illustrations for the words *kîɲ* (HL) ‘tomorrow’ and *sabûn* (L HL) ‘soap’, as shown in Figure 3.6 and 3.7, respectively. Figure 3.6 shows a simple fall while Figure 3.7 shows a complex fall (rise-fall) on the second syllable.

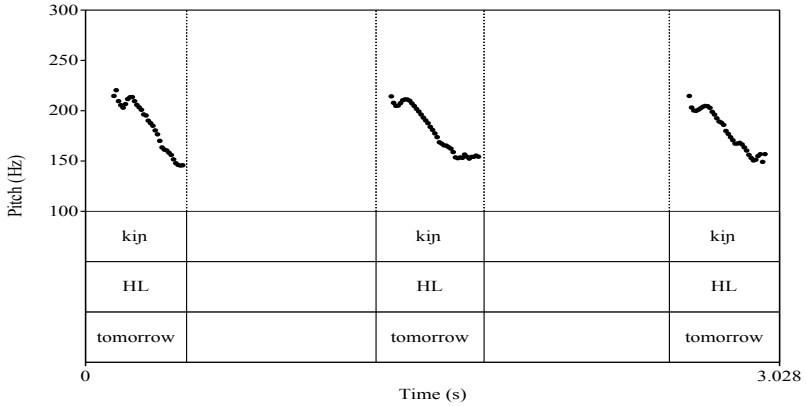


Figure 3.6 Three repetitions of the word *kîɲ* showing a simple falling contour

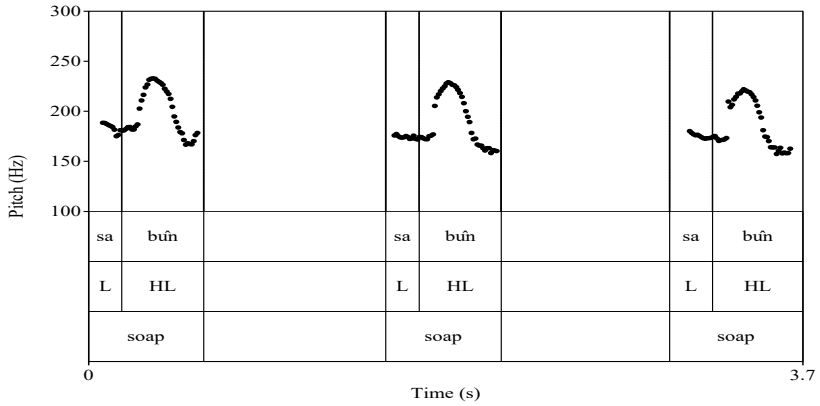


Figure 3.7 Three repetitions of the word *sabûn* showing complex (rise-fall) contour

One of the questions that might be of interest here is the difference between the fall noted in Figure 3.1, which is attributed to low tone, and that in Figure 3.7 attributed to a simple falling contour, for the words *pi* 'water' and *kîɲ* 'tomorrow', respectively. Based on the pitch values for these particular words it is noted that the difference between the maximum and the minimum pitch is 35.30Hz for *pi* and 67.52Hz for *kîɲ.* This indicates that *kîɲ* (HL) has a phonetically significant fall than *pi* (L).

For the case of a rising contour, two monosyllabic words *lăw* ‘cloth’ and *cwă* ‘tamarind’ are used as examples for a simple and a complex rise, respectively. In Figure 3.8, the word *lăw* shows a rise from a level low tone while in Figure 3.9 the word *cwă* begins with a fall from high pitch (i.e. 220Hz) targeting a low tone below 200Hz, which is then followed by a shallow rise, hence a complex rise (fall-rise).

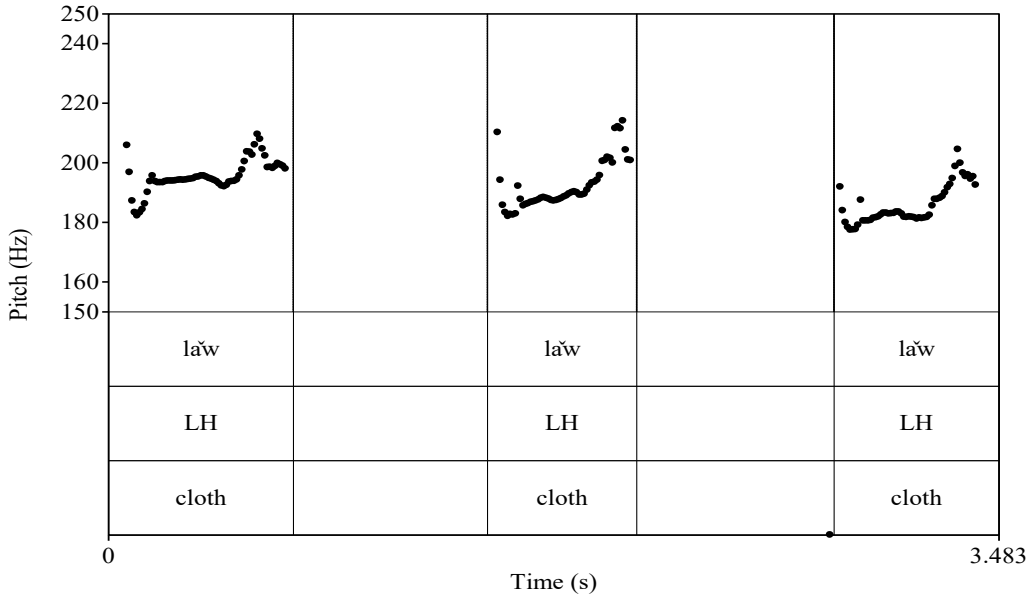


Figure 3.8 Three repetitions of the word *lăw* ‘cloth’ showing a simple rise

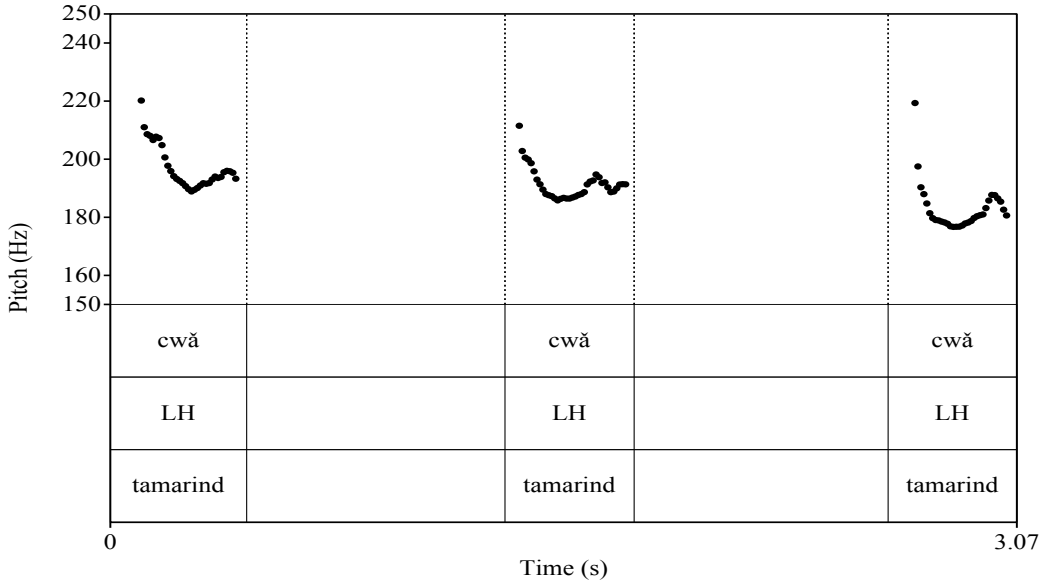


Figure 3.9 Three repetitions of the word *cwă* 'tamarind' showing a complex rise (fall-rise) contour

Generally, in Luo, HL is more productive and widely used than LH as it is found in many words including a number of loanwords such as *sabûn* 'soap' and *sûm* 'poison' which are borrowed from the Swahili words *sabuni* and *sumu,* respectively. On the other hand, a rising tone is less productive as it is found only in a few words as an underlyingly lexical tone. The rising contours are mostly found in derived contexts, which will be discussed later in section 3.3. It is not uncommon that languages prefer falling to rising contours, but it can also imply that the rising contour is no longer productive in the language. Regarding the environment for contours, languages which consider a syllable as a tone bearing unit normally have specific rules on contour occurrence. For example, Mende (a Niger-Congo language of the Mande branch spoken in Sierra-Leone and Liberia) allows contours only on word-final syllables (Leben, 1973), while Kunama (a Nilo-Saharan language spoken in Eritrea and Ethiopia) allows contour tones medially but only on heavy syllables (Connell et al, 2000). Luo seems

to allow contour tones on both light and heavy syllables, that is monomoraic and bimoraic syllables. However, it is argued that phonetically light syllables tend to be lengthened when they have a contour tone (Yip, 2002).

Based on the examples illustrated above, it can be concluded that Luo has two types of falling contours, HL and LH. This is in line with the autosegmental-metrical theory which also advocates for a single contour (HL or LH) as a phonological abstraction that captures all the other configurations. The various types of contours such as high falling, extra high falling, long high falling, falling-rising and rising-falling, as analysed by Tucker (1994), are just phonetic variations with respect to other orthogonal factors. To sum up on Luo lexical tones, the language has two underlying level tones, H and L, which can be realized differently in different contexts. The two tones form two types of contours, namely falling and rising which can also appear underlyingly as lexical tones. This is contrary to Omondi (1982) who argues that contours are only derived.

3.1.3 Grammatical tones

Besides lexical function, tone in Luo also has a grammatical function. This function can be categorized into two: to show aspectual distinctions or word class distinctions. The former is portrayed when particular tones are associated with specific tense/aspect/mood (i.e TAM). Such tones are known as melodic tones. Odden and Bickmore (2014) point out that melodic tones which signal verb tense-aspect are obtained by modifying the tone(s) of the root or

stem. In Luo, for instance, the underlying tonal pattern for verbs is L, as in *maŋo*[8] 'to find'. The progressive aspect is marked by inserting the grammatical or melodic H tones which replace the underlying tones. Odden (1996) proposes an analysis for this that verbs can be assigned a floating H which is mapped to the first or second stem vowel as determined by tense-aspect of the verb. For the case of Luo progressive aspect, it can be argued that the first vowel receives a H tone, which then spreads unboundedly to the rest of the verb when it is non-final as in (31). If the verb is sentence final as in (32) high tone spreading to the final syllable is blocked. The lack of high tone spreading to the final syllable indicates the presence of a low boundary at the end of declarative sentences. This is not only observed in Luo but in other languages as well. Melodic Hs remain the same even when the tense changes from present to past progressive as in (33). The only difference noted here is the past tense being marked by the prefix *né-*.

31. ó-máŋó náŋgá
    SM3SG-PROG.find cloth
    'S/he is looking for a cloth.'

32. ó-máŋò
    SM3SG-PROG.find
    'S/he is looking for (something).'

[8] Luo verbs are basically disyllabic (also referred to as full stem verbs) in their base form. The vowel sequences -ie-/-uo-/-ue- as in *riembo* 'to expel', *luoro* 'to fear' and *yueyo* 'to rest' are considered diphthongs and thus belong to the same syllable (see Tucker, 1994 chapter seventeen to twenty one for a detailed analysis of verbs in the language).

33. né-ó-máɲó náŋgá
PST- SM3SG-PROG.find cloth
'S/he was looking for a cloth.'

The perfective or completive aspect, on the other hand, is realized with a H tone inserted on the first syllable of the stem or root, which in turn surfaces with a falling contour (HL) as exemplified in (34). This is contrary to Tucker (1994: 354) who argues that the basic tonal pattern for perfective or completive aspect is L throughout the verb. In addition to this pattern, the perfective aspect is also marked by a prefix *o-* (which is homophonous with the 3rd person singular *o-*).

34. o-mâɲo náŋgá
SM3SG-PRF.find cloth
'S/he has looked for a cloth.'

Additionally, tone in Luo can also have a derivational function, by changing words from one class to another. For example, the low-toned words *ber* 'good' and *ler* 'clean/bright' in (35a) and (35c) respectively are adjectives but the same words when are high-toned as in (35b) and (35d) they change to nouns.

35. (a) pala-ni ber
knife-DEM good
'This knife is good.'

(b) bér mar pala

goodness of knife

'The goodness/beauty of a knife'

(c) ot ler

house clean/bright

'The house is clean/bright.'

(d) lér mar Yesu

brightness of Jesus

'The cleanliness/brightness of Jesus'

3.2 Tonal patterns

This section presents the possible tonal patterns in Luo. Firstly, monosyllabic words exhibit four patterns reflecting the lexical tones discussed above. Examples are given in (36) below.

36. pi L 'water'

món H 'women'

cwă LH 'tamarind'

sûm HL 'poison'

Disyllabic words have six patterns, as illustrated in (37).

37. mbaka L.L 'conversation'

nonó L.H 'empty'

sabûn L.HL 'soap'

ɲákó	H.H	‘girl’
ɲíen	H.L	‘new’
mácôn	H.HL	‘old’

Multisyllabic words (with three syllables or more) also have six possible patterns, as illustrated in (38).

38.

ɲakwaro	L.L.L	‘grandchild’
odumá	L.L.H	‘maize’
acíel	L.H.L	‘one’
obámblá	L.H.H	‘dried fish’
ríémbô	H.H.HL	‘to drive’
omâɲo	L.HL.L	‘he has found’

The tonal patterns presented above are either lexical or grammatical tones specified on the syllable as the TBU. H and L can occur anywhere in a syllable while LH and HL are restricted to syllable finally or at the penult. As noted earlier above, a falling contour is preferred to a rising contour. Nouns exhibit more varied patterns than verbs and adjectives. Adjectives mainly have L or HL patterns. Tucker (1994) notes that the H tone in adjectives is avoided because it tends to change an adjective into a noun.

3.3 Tonal processes

3.3.1 Tonal contraction

Tonal contraction occurs when a segment is deleted and its tone is carried by the rest of the word. In Luo, Omondi (1982) points out that tonal contraction can occur when the final vowel in a full imperative verb is deleted to form a short imperative. She posits that a full imperative verb has the final H toned morpheme *-í.* When this morpheme is deleted to form a short imperative, its tone becomes floating and is carried by the first syllable which is low-toned and in turn creating a rising contour as shown in (39)[9].

39.

Verb	**Full imperative**	**Short imperative**	
nindo	nindí	nǐnd	'sleep'
riŋgo	riŋgí	rǐŋg	'run'
tedo	tedí	těd	'cook'

3.3.2 High tone spreading

High tone spreading (henceforth HTS) is another tonal process in Luo tonology. HTS is treated as tonal assimilation in Omondi (1982) and Tucker (1994). In this process the high tone spreads rightwards targeting a L-toned or toneless syllable. When H-tone spreads to a low-toned word it does not override it but instead surfaces as a falling contour as shown in (40) and (41). In these examples we see the low-toned adjective *liet* 'hot' and object *le* 'axe' surface as HL.

[9] The domain of use and motivation for the difference between full and short imperative is not very clear as some verbs have both full and short forms while others do not.

40. pálá liet → [pálá líèt]
ochre hot
'Ochre is hot.'

41. ɲi nígí le → [ɲi nígí lê]
'People have axe.'

Tucker (1994: 338)

H-tone can also spread to a toneless syllable making it surface with a H-tone as seen in (42). In this example, we see the H-tone of the subject *món* 'women' spreading to the first syllable (which is toneless) of the verb *otêdo* 'have cooked'. It should be noted that subject prefixes are underlyingly toneless and tend to receive a H or L depending on the TAM tonal specifications.

42. món o-têdo → [món ótêdo]
women PRF-cook
'The women have cooked.'

In the examples given above, it is observed that HTS can occur between a subject and an adjective (40), a verb and an object (41) and a subject and a verb (42). This suggests that there are no restrictions to the boundaries that can be crossed. Secondly, it is observed that HTS between constituents in Luo is unary and non-iterative. This is because it only spreads once to the following syllable. This kind of spreading has been identified as High Doubling in Bantu (see, for example, Bickmore & Kula 2013).

3.3.3 Downstep

Downstep is the prominent tonal process in Luo. Although an in-depth analysis of downstep in Luo particularly at the level of a sentence is found in chapter five, in this section I will highlight the environments where downstep occurs as presented in the previous literature on the language. Tucker and Creider (1975) and Tucker (1994) are the works with an extensive study of downstep in Luo. In these works, Tucker and Creider argue that Luo has four types of downstep, namely explicit, implicit, grammatical and built-in downstep. Explicit downstep occurs when a low tone precedes a sequence of high tones. Here, a phrase or a sentence enters a lower key[10] and all the subsequent highs (including high falling tones) are also lowered. Tucker shows in (43) that the first high tone of the word *món* 'women' which spreads to the prefix *o-* begins higher at 40[11], but after the introduction of a low tone on the verb *owito* the subsequent highs are lowered to 25. Meanwhile, the final high is also lowered to 17.

43. món ówito ꜜnáŋgá móró oꜜkó

    25.40:40:25\20:14:25:25:25:25:15\7:17

    women throw cloth other outside

    'The women have thrown the other cloth away.'

    (Tucker, 1994:342)

Yip (2002: 148) considers the phenomenon in (43) as downdrift rather than downstep and distinguishes it as follows: Downstep is the lowering of a high tone in the absence of a

---

[10] For convenience, I have used an exclamation mark (!) to mark downstep instead of an apostrophe used by Tucker.

[11] Note that these are musical values known as lamellae obtained from tonometer readings as described in section 4.1. The highest value is 56 (Tucker, 1994:323)

surface low tone, while downdrift is lowering of high tone after an overt low tone. Gussenhoven (2004) considers the former as automatic downstep and the latter as non-automatic downstep. Thus, in Yip's and Gussenhoven's view, Tucker's explicit downstep in (43) is downdrift or automatic downstep, respectively, because of the preceding overt low tone.

Moreover, the sentence in (43) shows that it is not only the immediate high which is realized lower but all the subsequent highs. And this is clearly observed when the initial high pitch was 40 and then the following highs lowered to 25. As argued by Gussenhoven (2004), when automatic downstep takes place any other following high tone will maximally reach the pitch of the preceding H. As (43) shows, the final high goes even lower to 17 pitch value. This could be due to the introduction of another L-tone in the word *okó* 'outside' which makes the high even lower compared to the previous highs. It can also be an effect of final lowering, which is an extra dose of lowering at the end of an utterance regardless of the preceding context (Gussenhoven 2004).

The second type of downstep is implicit downstep. Here the lowering of H tones is caused by a covert L-tone. Yip (2002) refers to this kind of downstep as real downstep while Gussenhoven (2004) considers it as non-automatic downstep. Tucker and Creider (1975) and Tucker (1994) report that in this type of downstep the falling tone is realized as high within a chain of speech but eventually a following high is lowered as if the falling tone were still present. Consider the examples in (44) below.

44. (a) or kéló ˈgúén

20:30:30:21-16

brother-in-law bring chickens

‘The brother-in-law brings chickens.’

(b) ruoθ dwá ˈpálá

20:24/33:17:17

chief want ochre

‘The chief wants ochre.’

(c) ruoθ dwá ˈsûm

22:27/35:31\5

chief want poison

‘The chief wants poison.’

(Tucker, 1994:343)

The pitch values of the objects in (44a-c) are lowered immediately after the verbs. According to Tucker (1994) there is a floating L which triggers the lowering. Again, as noted earlier in (43), all these words are sentence final. Therefore, their lowering might be triggered by an intonational tone superimposed on the final syllable or a string of syllables as a marker of finality. In this case, it is not necessarily that there is a floating L tone which triggers downstep. Having a downstepped H sentence finally is a form of grammaticalizing final lowering (Gussenhoven, 2004).

The third type of downstep is grammatical downstep. Here the verb is said to be in a downstep relation with its subject regardless of whether the subject is high toned or low toned. And this phenomenon is only noted in the imperfective verb which has a H H tonal sequence. The perfective verb does not show this kind of relationship (Tucker 1994:343). Compare (43) and (45) with perfective and imperfective verbs, respectively.

45. món ˡwító ˡnáŋgá móró okó
    27/40:31:31:24:24:24:24:14\9:17
    women throw cloth other outside
    'The women are throwing the other cloth away.'

Regarding the grammatical downstep in (45) attention is paid to the subject and its immediate verb. Other forms of downstep seen are already identified above as either implicit or explicit or as automatic or non-automatic downstep. The most intriguing thing in grammatical downstep is that there is no obvious evidence for low tone triggering downstep. It seems this form of downstep relies on the fact that whenever there are two adjacent highs then the second is downstepped. This phenomenon is also observed in other languages. For example, Kula and Hamann (2016) in their analysis of downstep in Bemba point out that there are cases where downstep is considered a reflex to remedy an OCP violation, a rule which prohibits identical adjacent features, in this case two high tones.

The last type of downstep is built-in or lexical downstep, as already introduced in section 3.1.1. This is found within words whereby the second high in a word with two adjacent highs is lowered. The list in (46a) shows downstep within nouns while (46b) shows downstep within verbs.

46. (a) ndé$^{!}$dé ‘purse’

ló$^{!}$só ‘speech’

(b) bí$^{!}$ló ‘to do the tasting’

kwá$^{!}$nó ‘to do the counting’

ó$^{!}$ló ‘to do the pouring’

In (46) downstep is observed without an overt or covert L tone trigger. For the case of nouns in (46a) there is no evidence for the spread of high to the following syllable and Tucker (1994) argues for both HH and H$^{!}$H as permissible patterns in Luo nouns. The same patterns are observed by Odden (1996: 462) in Kishambaa nouns *ŋgó$^{!}$tó* ‘sheep’ and *nyó$^{!}$ká* ‘snake’. As already argued, this can be adherence to OCP which prohibits adjacent Hs. On the other hand, the qualitative verbs[12] in (46b), which surface with H$^{!}$H tonal pattern, are underlyingly L-toned when in their transitive forms, suggesting that the H$^{!}$H pattern has a grammatical role.

[12] These are verbs which cannot take an object despite describing actions associated with an object (Tucker, 1994:413).

3.4 Tonal typology

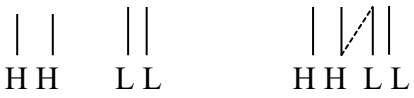
Tonal typology is one of the longstanding interest of theorists and fieldwork researchers (Hyman, 2010). The main concern has been to typologize a language based on which tones are phonologically specified or underspecified. The proposed systems are equipollent and privative (Hyman, 2001, 2010). The former includes those languages which phonologically specifies both H and L tones (H vs L) and the latter includes those languages where H is marked and L is unmarked (H vs ∅). In this system, only H-tone is specified and any other syllable receives a default L-tone or remains toneless, a system which has been proposed for Bantu.

Hyman (2001) proposes some potential criteria to establish whether both High and Low tones are phonologically activated in a language. The first one is that if a language allows either a rising LH or falling HL contour on a single tone bearing unit, be it a syllable or a mora, as illustrated in (47), then the possibility is that both H and L are active. Note that in Luo the syllable is the TBU and there are numerous examples of rising and falling contours on a single syllable as in *cwă* 'tamarind', *lăw* 'cloth', *kîɲ* 'tomorrow' and *sabûn* 'soap'. This criterion includes even the derived contours from processes such as tone spreading as observed in section 3.3.2 where the H-tone of the noun *pálá* 'ochre' spreads to the low-toned word *liet* 'hot' resulting in a falling contour *líèt.* This is further illustrated in (48).

47.

σ	σ	μ	μ
⋀	⋀	⋀	⋀
HL	LH	HL	LH

48. pálá + liet → pálá líèt ‘Ochre is hot.’



The second proposed criterion is that a H- or L-tone should occur as a floating tone, and in some cases, motivates or blocks certain phonological processes. This has also been observed in Luo through tonal contraction in section 3.3.1. In tonal contraction a vowel is deleted, and its tone becomes floating and carried by another syllable. An example for this is given in (49). In this example, the vowel of the full imperative *-í* is deleted in contracted forms and its high tone carried by the rest of the syllable. This process suggests that the H-tone is phonologically active, but it does not override the underlying L-tone. Thus, the floating H-tone and the L-tone both surface as a rising contour.

49. nindí ⟶ nǐnd ‘sleep’

In this respect, further evidence is given by Tucker (1994) when presenting the implicit downstep case. In this case, a floating L causes downstep of the following H (see section 3.3.3 for more details). An example is given in (50). Tucker argues that the verb *kélô* ‘to bring’ has an underlying falling contour (HL) on the second syllable and the L-tone of the contour dowsteps the subsequent H in the following word *gúén* ‘chicken’. Here, we see the floating L tone motivating the downstep of H.

50. or kélô gúén ⟶ [or kéló ꜜgúén]

in-law bring chicken

‘The in-law brings chicken.’

Based on the examples given, it is evident that both L and H tones are phonologically activated in the language, suggestive of an equipollent system (H vs L). Nevertheless, not all L-tones in the language are specified. We have noted incidences as that in (42) where a subject prefix is toneless. It can therefore be argued that Luo is (H vs L vs ∅). Hyman (2001) proposes that a language may start out as a privative (H vs ∅) system where one of the tones is unspecified but become equipollent where both tones are active during derivation.

Chapter Four

Research Methodology

4.0 Introduction

This chapter presents the methodology that was used in collecting and analyzing the data of this study. Section 4.1 provides information on the research area and the context in which the data were collected. Section 4.2 elaborates on the participants who took part in this study. The materials designed to obtain data are presented in section 4.3. The recording procedure, data processing and analysis procedure are presented in section 4.4 and 4.5. Lastly, the chapter ends by highlighting the challenges and limitations faced while conducting the research and points out the ethical issues observed.

4.1 The research area and fieldwork

Data were collected from Rorya and Tarime districts in the Mara region which is found in Northern Tanzania around Lake Victoria and the Kenyan border. Rorya district is the predominantly Luo speaking area in Tanzania. Tarime district in the Eastern parts of the region is predominantly a Kuria speaking area but with a significant number of Luo speakers, especially in the urban area. More information about Luo speaking areas in Tanzania is presented in chapter 1 section 1.1. The first fieldwork for this research was conducted in July and August 2017 and the second one in September 2018. With participants drawn from both areas, there was a balance between those who claim to be pure Luo and those who claim Bantu origins (see Kihore, 2005).

4.2 The participants

Four native speakers, two females and two males, participated in this study. All speakers declared Luo as their first language and Swahili, which is the national and first official language in Tanzania, as their second language. All the participants declared some knowledge of English but none declared to speak any other Bantu language. Their ages range between 18 and 55. Among the participants, three were born and raised in Rorya district. One participant relocated to Tarime to work as a primary school teacher. The other two are still living in Rorya – one is a secondary school student and the other one a primary school teacher. One participant was born and raised in Kenya but has been living in Tarime for more than twenty years working as a tailor.

There was in addition, another three participants, two males and one female, who are elderly and in their 60s, 70s and 80s. These were specifically interviewed for historical background on the Luo people and understanding of the social setup of the Luo speaking areas. All of them were born and raised in Rorya district and speak Luo as their first language. Two participants are from Luo Imbo division, an area known to be occupied by pure Luo speakers. The other one is from Nyancha division and she claims to be originally a Rieri but has grown up speaking Luo as her first language.

Participants were approached individually after being recommended by others and asked to take part in the study. The factors considered were that the participant is a native speaker, uses Luo in most of his or her daily activities, and is ready and available to take part and be recorded.

4.3 Test materials

The materials designed to obtain data for analysing Luo intonation comprise of varied sentences. The materials are largely based on scripted Luo sentences to be read and also a few Swahili sentences to be translated into Luo. Additionally, there were picture-based tasks meant to trigger spontaneous speech to complement the scripted sentences. These test materials are explained below.

4.3.1 Sentence lists

The sentences comprised of simple declaratives, questions, complex sentences, focused constructions, dislocations, topics, and clefts. There were five sets of sentences labelled A, B, C, D and E. Each sentence type had a separate set of materials.

Set A: The material presented in this set served as a baseline to investigate the intonation of simple declarative sentences. In order to evaluate F0 effects for declarative sentences both syntactic and phonological compositions of sentences were observed. Basically, Luo is an SVO language. Therefore, the sentences were made up of the following constituents: NP + VP + (PP) + (AdjP). Some sentences had one object (i.e. direct) as in (51a) or two objects (i.e. direct and indirect) as in (51e) and (51f). Other sentences had prepositional phrases (see 51b, c, d and g) as well as adverbs of place (see 51h). Having different syntactic constituents was meant specifically to investigate the relationship between syntactic constituencies and prosody. Constituents were also controlled for tone to ensure different combinations of tone so as to test whether this had an effect on intonation.

Moreover, in order to ensure that the aspectual distinction of the verb is captured, every sentence had a perfective and its imperfective counterpart. It should be noted that, in Luo, the grammatical tones for imperfective aspect are H H H (HHHL sentence finally) pattern and L HL L pattern for the perfective aspect, compare (51b) and (51c) as well as (51e) and (51f). Given this nature of the verb, getting a declarative sentence with all-low tones was not possible. However, a nearly-all-low tone sentence was used. All other targeted words remained either low-toned or high-toned. Using same-tones or a sequence of alternating tones is a common technique to investigate downtrends particularly in tone languages (Jun and Fletcher, 2014).

To evaluate the F0 effects at the phrase edges, words with high, low, rising and falling tones were tested in the final and non-final position. In order to avoid segments disrupting F0 the use of less disruptive consonants (i.e. sonorants) was factored in. But this was not fully achieved because striking a balance between recommended segments with the required tonal targets and yet meaningful sentences was not possible for all the cases.

Concerning the length of the sentences, the number of syllables was also considered. The shortest sentence had four syllables while the longest had twelve syllables. Note that most words in Luo are either monosyllabic or disyllabic. There are only a few multisyllabic words since the language is not agglutinating. In total there were 132 declarative sentences each recorded three times making a total of 396 tokens. The analysis for these sentences is presented in chapter 5 section 5.1.

51. **Declarative sentences**

(a) ɟi cámó kuon

people PROG.eat ugali

'People are eating ugali.'

(b) món tédó réc gí yién

women PROG.cook fish with firewood

'The women are cooking fish with firewood.'

(c) món otêdo réc gí yién

women PRF.cook fish with firewood

'The women have cooked fish with firewood.'

(d) ɟi míélό θum gí kwac

people PROG.dance music with leopard

'People are dancing music with a leopard.'

(e) ɲiθíndo míyó ɟadúoŋ mo-tí sabûn

children PROG.give man REL.old soap

'The children are giving old-man soap.'

(f) dayo omîyo wuoro ɲuka

grandmother PRF.give father porridge

'The grandmother has given father porridge.'

(g) ruoθ ði e ot

chief PROG.go in house

'The chief is going in the house.'

(h) ɟapúóɲɟ ríémbó ndígá okó

teacher PROG.ride bicycle outside

'The teacher is riding a bicycle outside.'

Set B: The material presented in this set served as a baseline to investigate the intonation of yes-no (polar) questions. Syntactically, yes-no questions are like declarative sentences except that yes-no questions have an optional particle '*be*' added at the beginning of the sentence as shown in (52a - c). Basically, some declaratives in set A were repeated in set B but with an added question particle. In total there were 54 yes-no questions which were produced three times amounting to 162 tokens per participant. The analysis of this sentence type is presented in chapter 5 section 5.2.1.

52. **Yes-No questions**

(a) be ɟi cámó kuon

Q people PROG.eat ugali

'Are the people eating ugali?'

(b) be món tédó réc gí yíén

Q women PROG.cook fish with firewood*'*

'Are the women cooking fish with firewood?'

(c) be món otêdo réc gí yíén

Q women PRF.cook fish with firewood

'Have the women cooked fish with firewood?'

Set C: In this set sentences were designed to investigate the intonation of wh-questions or content questions. In Luo, question words are realized in-situ. The questions that were tested are those with the following question words: *ŋa* 'who' as a subject, shown in (53a); *ŋa* 'who' as an object, shown in (53b); *aŋó* 'what' as in (53e to 53g); *káŋé* 'where' as shown in (53c); and *karaŋó* 'when' in (53d). A sentence with multiple question words was also included as in (53e). Although Luo is a question in-situ language, question words can be fronted as in (53f). The syntactic and phonological compositions of the content questions in this set are derived from declarative sentences in set A. Therefore, the syntactic constituents, tonal specifications and number of syllables are the same as their declarative counterparts. In total there were 54 tokens (18 sentences x 3) per participant. The analysis for content questions is presented in chapter 5 section 5.2.2.

53. **Content questions**

(a) **ŋa** má ŋíéwó obámblá

who REL PROG.buy dried fish

'Who is buying dried fish?'

(b) dayo míyó **ŋa** ɲuka

grandmother PROG.give who porridge

'Who is grandmother giving porridge?'

(c) ɟapúóŋɟ ríémbó ndígá **káɲé**

teacher PROG.ride bicycle where

'Where is the teacher riding a bicycle?'

(d) ɟapúóŋɟ lémó **karaŋó**

teacher PROG.pray when

'When is the teacher praying?'

(e) **ŋa** má lwókó **aŋó** **káɲê**

who REL PROG.wash what where

'Who is washing what where?'

(f) **aŋó** má í-dwárô

what REL 3SG-want

'What is it that you want?'

(g) ɟapur bíró gi **aŋó**

farmer PROG.come with what

'Whatis the farmer coming with?'

Set D: The materials in this set were designed specifically to investigate the F0 effects on 'information structure'. Güldemann et al. (2015) define information structure as how speakers structurally encode propositional content with respect to their assessment of knowledge that is shared (or not) by the interlocutors in a communicative situation. Information structure includes givenness, focus and topic.

Focus can be obtained in-situ as well as ex-situ. The most common one is an answer to a question, which is also referred to as in-situ focus. Secondly, focus can be obtained from syntactic manipulation of sentences such as dislocations and cleft constructions (Güldemann et al., 2015). This form of focus is also referred to as ex-situ focus. Thus, to obtain in-situ focus I used a different technique which is discussed in section 4.3.2. For ex-situ focus as well as topics the sentences in (54), (55) and (56) were used. These include topicalized subjects, left dislocations or topicalized objects, right dislocations and clefts. Their analysis is presented in chapter 6. In this set, however, materials were in Swahili (a language declared to be a second language by all participants). The participants were then asked to translate the

Swahili sentences into Luo, after the context was made clear. Translation technique was used because Swahili has an overt pronoun which makes it easier to identify a topicalized constituent.

The sentences in (54) were used to elicit data on subjects as topics. The respective subject markers are marked in the verb but also preceded by a nominal subject to create a topicalized sense. Thus, the subject is a topic and the VP a complement. In total there were 5 sentences which were produced three times each. This makes a total of 60 tokens (i.e 5 sentences x 3 repetitions x 4 participants). The analysis for topicalized subjects is presented in chapter 6 section 6.4.3.

54. **Topicalized subjects**

(a) Mimi, ni-na-pend-a wa-toto
I SM1SG-PRS-love-FV 2-children
'I, I love children.'

(b) Sisi, tu-na-pend-a wa-toto
We SM1PL-PRS-love-FV 2-children
'We, we love children.'

(c) Mama, a-ta-kwend-a soko-ni kesho
mother SM3SG-FUT-go-FV market-LOC tomorrow
'Mother, she will go to the market tomorrow.'

(d) Watoto, wa-ta-chez-a kwenye m-changa kesho mchana
children, SM3PL-FUT-play-FV in 3-sand tomorrow afternoon
'The children, they will play in sand tomorrow afternoon.'

(e) Baba, a-ta-kwend-a ku-vua samaki jioni
Father, SM3SG-FUT-go-FV INF-catch fish evening
'Father, he will go to catch fish in the evening.'

Regarding topicalized objects, also referred to as left dislocations, as shown in (55), the objects have their relevant markers on the verb. Note that, when there are two objects like in (55d), only one object can be marked in the verb, and in this case, it is the indirect object 'people'. For topicalized objects there were 4 sentences as outlined in (55). This amounts to 48 tokens (i.e. 4 sentences x 3 repetitions x 4 participants). The analysis for topicalized objects is presented in chapter 6 section 6.4.2.

55. **Topicalized objects/left dislocations**

(a) Chui, wa-toto wa-na-mu-u-a
Leopard, 2-children SM3PL-PRS-OM3SG-kill-FV
'Leopard, the children are killing it.'

(b) Uji, m-pishi a-na-u-korog-a na m-wiko jiko-ni
Porridge, 1-cook SM1SG-PRS-OM1SG-stir-FV with 3-stick kitchen-LOC
'Porridge, the cook is stirring it with a cooking stick in the kitchen.'

(c) Wa-toto, mama a-me-wa-tum-a soko-ni
2-Children, mother SM3SG-PRF-OM3PL-send-FV market-LOC
'Children, mother has sent them to the market.'

(d) Ugali, baba a-na-wa-pat-i-a wa-tu
Ugali, father SM3SG-PRS-OM3SG-give-APPL-FV 2-people
'Ugali, father is giving it to people.'

Apart from left dislocations, right dislocations were also investigated. For the case of right dislocations the assumption is that the hearer seems not to understand the person or thing being referred to, the speaker is then prompted to mention that person or thing as an afterthought. This means both subjects and objects can be right-dislocated. In order to achieve this the dislocated constituent should have an antecedent in the previous clause. Here the nominal subject or object co-ccurs with its respective marker on the verb. In this set (56a) and (56b) were used as examples of sentences with dislocated subjects and (56c) and (56d) as examples of dislocated objects. There was a total of 48 tokens (i.e 4 sentences x 3 repetitions x 4 participants). The analysis for right dislocations is presented in chapter 6 section 6.4.1.

56. **Right dislocations**

(a) A-na-wa-pat-i-a wa-tu ugali, baba
SM3SG-PRS-OM3SG-give-APPL-FV 2-people ugali father
'He is giving people ugali, father.'

(b) Wa-me-pik-a ugali haraka, wa-nawake
SM3PL-PRF-cook-FV ugali quickly 2-women
'They have cooked ugali quickly, the women.'

(c) Wa-nawake wa-me-u-pik-a vizuri, uji
2-Women SM3PL-PRF-OM3SG-cook-FV nicely porridge
'The women have cooked it nicely, porridge.'

(d) Wa-toto wa-me-zi-nunu-a, samaki
2-Children SM3PL-PRF-OM3PL-buy-FV fish
'The children have bought them, fish.'

Another form of focus that was investigated is clefts. The cleft constructions used are indicated in (57a-d). In these examples, the first constituent is focused. This sub-set also had 48 tokens (i.e. 4 sentences x 3 repetitions x 4 participants). The analysis for cleft constructions is presented in chapter 6 section 6.3.

57. **Cleft constructions**

(a) M-toto ndiye a-na-end-esh-a baiskeli
1-child 1.REL SM3SG-PRS-move-CAUS-FV bicycle
'The child is who rides the bicycle.'

(b) Maji ndi-yo m-toto a-na-kunyw-a

9.water 9.REL 1-child SM3SG-PRS-drink-FV

'Water is what the child drinks.'

(c) Kijiji-ni ndipo pa-na-end-esh-w-a baiskeli

village-LOC 16.REL SM3SG-PRS-move-CAUS-PASS-FV bicycle

'Village is where the bicycle is ridden.'

(d) Baba ndiye a-na-ye-wez-a ku-tu-omb-e-a

father 1.REL SM3SG-PRS-REL-can-FV INF-OM1PL-pray-APPL-FV

'Father is who can pray for us.'

Set E: The materials in this set were meant to investigate the intonation of complex sentences as well as the higher levels of phrasing and how they are cued. The sentences included subordinate, coordinate and relative clauses. As it was in set D, here participants were also required to translate the Swahili complex sentences into Luo. The analysis for complex sentences is presented in chapter 7. Examples (58a) to (58c) are sentences with subordinate clauses, (58d) and (58e) are examples of restrictive relative clauses, (58f) and (58g) are examples of non-restrictive relative clauses while (58h) and (58i) are examples of coordinated clauses. This part comprised 14 sentences produced 3 times by each participant. This makes a total 168 tokens.

58. **Complex sentences**

(a) Kama ku-na m-tu a-na-tak-a cha-kula na a-pik-e

if 15.there-is 1-person SM3SG-PRS-want-FV 7-food CONJ SM3SG-cook-FV

'If someone wants food, let him cook.'

(b) Ni-li-kuw-a ni-na-lim-a shamba l-angu a-li-po-fik-a

SM1SG-PST-was-FV SM1SG-PRS-cultivate-FV 5.farm 5-POSS SM3SG-PST-LOC-arrive-FV

'I was cultivating my farm when he arrived.'

(c) Kwasababu u-me-chelew-a si-ta-onge-a na wewe

because SM2SG-PRF-late-FV NEG.SM1SG-PST-talk-FV CONJ you

'Because you are late, I won't talk to you.'

(d) M-vulana a-li-ye-simam-a pembe-ni ya meza ni mw-anangu

1-boy SM3SG-PST-REL-stand-FV beside-LOC of 9.table COP 1-child

'The boy who is standing beside the table is my son/child.'

(e) M-tu a-li-ye-kuj-a ni kaka y-angu

1-person SM3SG-PST-REL-come-FV COP 9.brother 9.POSS

'The person who has come is my brother.'

(f) M-ke w-angu, ni-na-ye-m-pend-a sana, ni mw-alimu
1-wife 1.POSS SM1SG-PRS-REL-OM3SG-love-FV much COP 1-teacher
'My wife, whom I love dearly, is a teacher.'

(g) Mwalimu Juma, ambaye a-me-andik-a ki-tabu, ni m-kali sana
teacher Juma who SM3SG-PRF-write-FV 7.book COP 1-strict very
'Teacher Juma who has written a book is very strict.'

(h) Tu-li-kwend-a ku-m-tembel-e-a lakini ha-tu-ku-m-kut-a
SM1PL-PST-go-FV INF-OM1SG-visit-APPL-FV but NEG-SM1PL-INF-OM3SG-find-FV
'We went to visit him but he was not there.'

(i) Baba a-na-shon-a nguo na mama a-na-pik-a cha-kula
9.father SM1SG-PRS-sew-FV 9.cloth CONJ 9.mother SM3SG-PRS-cook-FV 7.food
'Father is sewing clothes and mother is cooking food.'

4.3.2 Picture-based tasks

There were two picture-based tasks. The first task aimed at triggering spontaneous speech to complement the scripted sentences outlined above in section 4.3.1. The second task aimed at eliciting focus. In the first task, participants were asked to talk about the things, the people and the activities that were depicted in the pictures. The main objective was for participants

to generate different types of sentences, particularly declarative and complex sentences from free speech. The themes in the pictures reflected the themes in the scripted sentences. There were 10 pictures with their themes outlined below.

Picture 1: A woman and a toddler sitting on the floor reading a book

Picture 2: A woman carrying a baby on her back while cooking at an open fire in a village setting

Picture 3: A couple of women selling fish and vegetables in a market

Picture 4: A woman riding a bicycle

Picture 5: Two people serving food to school children

Picture 6: Children washing clothes at the river

Picture 7: A man and a woman dancing in the rain

Picture 8: A man cooking outside the house and a little girl running

Picture 9: A young man eating ugali and beans in a classroom setting

Picture 10: A man, wearing traditional attire, is herding a flock of cows

The second task had three pictures with scripted questions. This task aimed to evoke narrow focus, alternative focus and corrective focus. This was done in a form of role play where one participant asked the other questions. The one who asked questions was given the pictures with scripted questions and the one who answered had pictures without the questions.

In order to capture narrow focus, in-situ questions were constructed as in (59a-e). These involved the subject being in focus as in (59a), the verb in (59b), the adverb in (59c) and the objects in (59d-e). The targeted words had low, high and falling tones.

59. **Narrow focus**

(a) má ŋa ni e pica

this who is in picture

'Who is this in the picture?'

(b) ó-tímó aŋó

SM3SG-do what

'What is he doing?'

(c) én káɲé

he where

'Where is he?'

(d) ó-rwákó aŋó

SM3SG-wear what

'What is he wearing?'

(e) ó-cámó aŋó

SM3SG-eat what

'What is he eating?'

Alternative focus was captured by the structures in (60) where the speaker was prompted to choose the right object, as depicted in the picture.

60. **Alternative focus**

(a) ó-cámó kuon gí réc koso gí ogânda
SM3SG-eat ugali with fish or with beans
'Is he eating ugali with fish or beans?'

(b) ó-tédó mchêle koso kuon
SM3SG-cook rice or ugali
'Is he cooking rice or ugali?'

(c) wuoro tímó aŋô to ɲaθî tímó aŋô
father PROG.do what CONJ child PROG.do what
'What is the father doing and what is the child doing?'

Corrective focus was elicited with yes-no question structures in 61. The speaker was required to give the correct information based on what he or she saw in the picture. The question in (61a) has its corresponding answer in (61b) while (61c) has its answer in (61d). The expected focus words are low-toned and high-toned, respectively.

61. **Corrective focus**

(a) (be) ó-cámó mchêle
Q SM3SG-eat rice
'Is he eating rice?'

(b) ó-cámó kuon

SM3SG-eat ugali

'He is eating UGALI'

(c) to gí-úsó maêmbe

Q SM3PL-sell mangoes

'Do they sell mangoes?'

(d) gí-úsó réc

SM3PL-sell fish

'They are selling FISH.'

4.4 Recording

Recording was done with an M-Audio MicroTrack II Professional Recorder with a Cliptec BMH699 microphone mounted on it. The wav files produced had a minimum recording rate of 44.1 kHz. The recording sessions for scripted sentences ranged between 20 and 40 minutes depending on the participants' convenience and availability for the session. Picture-based tasks were recorded between 5 and 10 minutes. After recording a set of similar sentences like declaratives or questions, participants took a break to relax and reconceptualise on the next task. As suggested by Jun and Fletcher (2014) breaks are important especially when recording a similar set of materials to avoid repetition effects that could occur in a single recording session. Scripted sentences were printed on paper whereby statements were marked by a full-stop (.) and questions by a question mark (?). Normally, tones are not marked in the

orthography, and even when they are marked most speakers do not understand them. Therefore, I had to explain the context to them before recording emphasizing on the aspectual differences which have tonal distinctions in the language. The sentences were read aloud three (sometimes four) times. For the picture-based tasks, coloured pictures were printed on paper and participants were asked to talk about the people, objects and activities depicted in the pictures.

4.5 Data processing

Audio wav files were transferred from a digital recorder to a computer. Thereafter, the pitch trends were computed by *Praat* software (Boersma, 2001). The acoustic analysis involved examining the fundamental frequency (F0) which is modelled in Hertz (herein referred to as Hz). The F0 is the physical correlate of pitch from speech waveforms (Pierrehumbert, 1980). Among the three (or four) repetitions the F0 contour of medial sentences was preferred. The determined pitch range floor was 75Hz. This was meant to capture even the lowest pitch ranges. In case of obvious faults, they were manually corrected by cutting them out. Syllables were manually labelled after listening to the sound and observing the formants and wave forms. Due to coarticulations and segment reductions it was sometimes difficult to accurately demarcate the syllable boundaries but overall labelling worked well. Tones (particularly lexical tones) were labelled for each syllable. Intonational tones were identified at their respective locations (e.g. at the edges). Pitch values were obtained from *praat* by selecting a vowel and then calculating the average pitch. Liberman and Pierrehumbert (1984) suggest that in order to reduce coarticulation effects, the measuring of pitch values should be at least

two frames from the consonantal release or closure.The following section gives further explanations of how data was analysed.

4.6 Data analysis

As noted above, intonation can be analyzed by observing the F0 contour. F0 can provide both qualitative and quantitative analyses. Qualitative analysis is based on the F0 contour shape which shows peaks, valleys and plateaus of a given utterance. Meanwhile, quantitative analysis is based on the phonetic description of speech by computing the quantitative values of the given tones. This kind of analysis is supported by Pierrehumbert (1980) who argues that rules can be explained in both quantitative features of F0 contour and the qualitative characteristics such as the F0 overall shape. Thus, the analyses in chapters 5 to 7 are mainly qualitative based on the F0 shape but at some points the pitch values were computed using Microsoft Excel. This was specifically done to measure the declination and final lowering effects.

According to Pierrehumbert (1980) F0 has some articulatory challenges, but it is still the most practical and accessible mode of analysis compared to other painstaking and difficult modes such as electromyographic studies of the laryngeal muscles and tracheal punctures. Another challenge when using F0 contour as a phonetic representation of intonation is the disruption caused by speech segments such as unvoiced sounds. Pierrehumbert (1980) notes that even voiced segments can have substantial effects on the F0. For example, F0 at the onset of the vowel after an unvoiced consonant is higher than a sonorant consonant. Also, there is a sharp fall near voiced obstruents and glottal stops. Moreover, high vowels tend to

raise the F0. Pierrehumbert notes that the difference between a high vowel and a low vowel can be 25 Hz at the same intonational context. Such effects make it difficult to identify what comes from intonation and what comes from segmental effects. To minimize these effects, scholars suggest maximum use of sonorants instead of obstruents or voiced sounds instead of voiceless sounds. This was, however, as we shall see in this work, difficult to avoid in some circumstances. This is because it is difficult to have a meaningful sentence with the targeted tones and yet have sonorants or voiced sounds only.

Theoretically, this study uses the Autosegmental-Metrical (AM) model to decide on tonal categories and intonational phonology as presented in Ladd (1996, 2008) as well as in Jun and Fletcher (2014). A detailed description of the theory is presented in chapter 2. In short, AM considers intonation as a sequence of tonal structures in both tonal and intonational languages. This tenet pays attention to the localized intonational events such as edge tones. Edge tones mark the left and right edges of a prosodic phrase. Furthermore, AM focuses on tonal targets rather than tonal movements or configurations. Here I consider the Hs and Ls as phonological abstractions which can be realized differently phonetically. This is important in African languages such as Luo which consider contour tones as sequences rather than units. Lastly, the theory recognizes global F0 trends. This helps in describing incidences of pitch range expansion or compression, which are global F0 effects.

4.7 Methodological challenges and limitations

All research is faced with some challenges and limitations. In this research, challenges range from participation, recording environment to data quality. First, during the first fieldwork one

male participant was unable to attend all the recording sessions because he had an emergency and had to travel.

Given that recordings were done at the participants' residences, we requested other residents to maintain silence to the maximum during recording sessions but there were external distractions which were beyond the control of the researcher. In some severe cases, like a motor vehicle passing, recording had to stop or sentences repeated because of the interruption. This disrupted the flow of the recording sessions and caused them to last longer than planned. Moreover, it was discovered during data processing that some data had faults like unexplained pitch jumps which when corrected affects the quality of the sentence. In such cases the data had to be abandoned making some cases inconclusive. Specific cases are highlighted in the respective chapters.

4.8 Ethical considerations

In line with good ethical practice, ethical approval from the University of Essex was granted as per the research regulations. In the field, participants were informed about the purpose of the research and what is expected of them before they agreed to participate. They were also informed of their rights such as the right to withdraw from the research, the right to ask questions and clarification at any point before, during and after data elicitation, the right to confidentiality and anonymity, as well as who to contact in case of any queries. These practicalities are recommended in Bell (2005) and Johnson and Christensen (2014). Before any recording was done, participants gave their informed consent in writing. In this work their initials instead of full names are used.

Chapter Five

Declaratives and Questions

5.0 Introduction

This chapter presents the intonation of declarative sentences and questions in Luo based on the Autosegmental-Metrical framework (Pierrehumbert 1980, Ladd 1996, 2008). As was earlier presented the theory was mainly developed to cater for pitch accents as the important tonal events in languages with sentence stress. For Luo as a tone language, I will mark intonational events based on the theory's use of global F0 trends which affect an entire utterance, such as, pitch register expansion or compression as well as localized phonological events such as pitch reset, downstep and final lowering. These global and local F0 trends have been previously used for other tonal languages (Ladd 1996, 2008). This chapter investigates the occurrence of these F0 features in simple declarative sentences and in simple questions.

For the notation, lexical tones will be marked as follows: H stands for high tone, L for low tone, HL for falling tone and LH for rising tone. For the case of intonational tones, I will follow the AM model notation, with diacritics added to differentiate them from lexical tones as follows: H% and L% will mark right high boundary and low boundary, respectively; -H will mark high register tone on the left edge; !H will mark a downstepped High at the intonational level.

The chapter is organized as follows: Section 5.1 discusses the intonation of declarative sentences with particular reference to simple sentences. Section 5.2 presents the intonation of

questions. Polar and content questions are discussed in sub-section 5.2.1 and 5.2.2, respectively. Finally, section 5.3 concludes the chapter.

5.1 Declarative sentences

Luo is an SVO language. A simple declarative sentence is made up of a subject followed by a verb which can carry one or two objects (i.e. direct and indirect object). The syntactic and phonological compositions of the declarative sentence samples discussed in this chapter are found in chapter 4 section 4.3.1. Generally, declarative sentences show a tremendous amount of downtrend in pitch contour which makes Luo a tone terracing language. According to Clements (1979) tone-terracing languages display regular register shift, a process which affects pitch realization of successive tones. This shift results in downward movement, notably downstep and declination. Tone-terracing is widely reported in African languages.

5.1.1 Downstep

Downstep is the most frequently reported form of terracing (Clements, 1979). It is shaped by the interaction between L-tones and H-tones. In this case, a H-tone is realized at a lower pitch relative to the preceding High. It is also possible to find that the downstepped H is at the same pitch level with the preceding L-tone or even lower. As discussed earlier, the domain of downstep differs from language to language but generally the lowering is caused by an intervening L-tone whether it is overtly or covertly present. When the lowering is triggered by an overt L, it is referred to as automatic downstep which is schematized as H L !H, but when there is no obvious L-tone trigger, then the lowering is referred to as non-automatic

downstep (Stewart, 1965). This is schematized as H (L) !H, with the brackets showing a floating L. In both cases, the downstepping environments are said to be local processes from adjacent tones (Connell, 2001).

Generally, in the literature, downtrends are termed differently and in most cases their meanings overlap. For example, Downing and Rialland (2016) use the term downdrift for automatic downstep and 'simply' downstep for non-automatic downstep. In the early works of Luo intonation by Tucker and Creider (1975) and Tucker (1994), automatic and non-automatic downstep roughly correspond to what they term explicit and implicit downstep, respectively. According to Tucker and Creider downdrift in Luo affects utterances with the same tones such as all-high or all-low tone utterances. In recent literature, this form of downtrend utterances is treated as declination. In this thesis, I will use the terms automatic and non-automatic downsteps as defined above. Their occurrence in Luo is further explored below.

5.1.1.1 Automatic downstep

As explained above, automatic downstep is caused by an associated L tone. This is observed in Luo declarative sentences with a sequence of Hs and intervening Ls. Consider sentences (62) and (63) below. Sentence (62) has a sequence of Hs with intervening Ls on the verb only. Meanwhile, sentence (63) which is relatively longer with twelve syllables has alternating Hs and Ls in every word except the verb, which has all Hs. Note that the falling contours on the verb *otêdo* in (62) and on the object *sabûn* in (63) are considered as a sequences of two levels, that is H and L.

62. H LHLL H H H

món otêdo réc gí yíén

women PRF.cook fish with firewood

'The women have cooked fish with firewood.'

63. L H L H H L HL L H L HL

ɲiθíndo míyó ɟadúoŋ mo-tí sabûn

children PROG.give man REL-old soap

'The children are giving the oldman soap.'

The F0 contour for (62) is presented in Figure 5.1. In 5.1(a), we see the H-tone on the first syllable *món* produced relatively high, approaching 250Hz. This H-tone spreads to the first syllable (*o*-) of the verb making its F0 higher despite being underlyingly L-toned. The H-tone spreading rule in Luo was discussed in chapter 3. Furthermore, there are intervening L-tones on the verb which subsequently lowers the following Hs. Arguably, there is a significant difference between the peak of a non-downstepped H-tone on the subject and the peaks of the subsequent downstepped Hs, which is approximately 50Hz. Once a downstep takes place, any following H-tone will maximally reach the pitch of the preceding downstepped one (Gussenhoven, 2004:100). This is what we observe on the final three syllables where H3 is lowered creating a ceiling for H4 and H5.

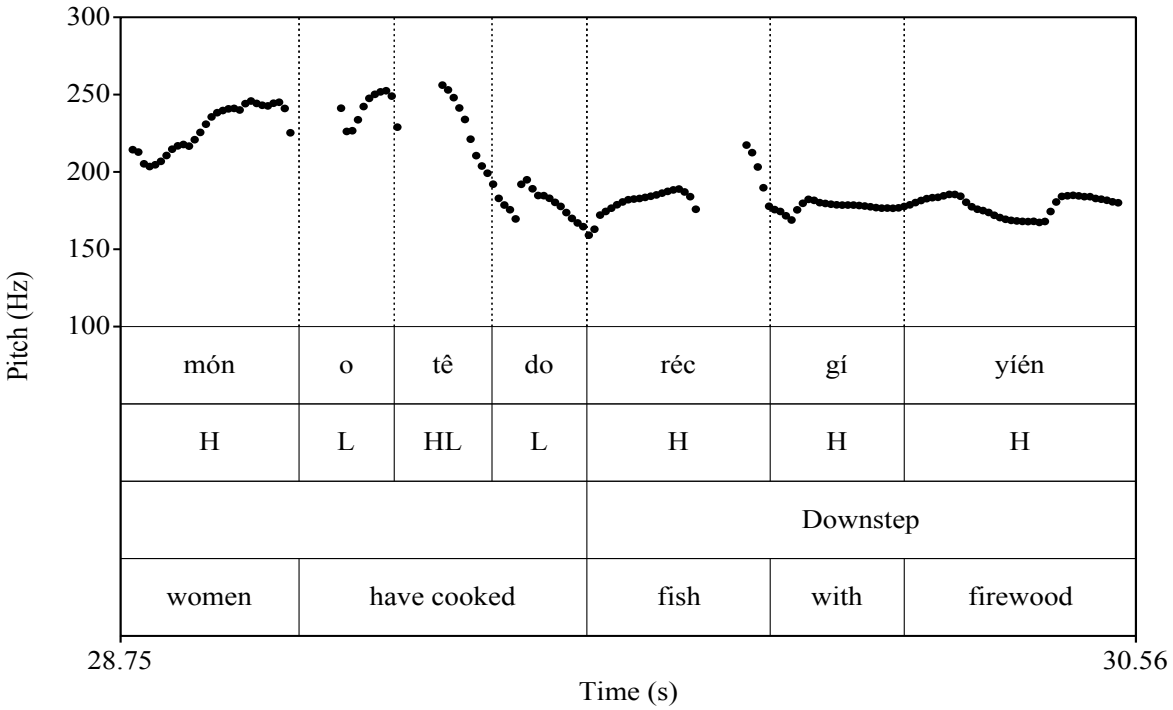


Figure 5.1(a): Pitch contour (in Hz) for sentence (62) showing downstepped Hs after an L-tone intervention; produced by a female speaker MG

Figure 5.1(b) reproduces the downstep effect already presented in 5.1(a) by showing the F0 values from four different speakers (2 females and 2 males) while 5.1(c) shows the average F0 values aggregated from the four speakers. In 5.1(b) speaker MA portrays a lower pitch range compared to the other three speakers and speaker SY exhibits an early fall in pitch noted in the second syllable.

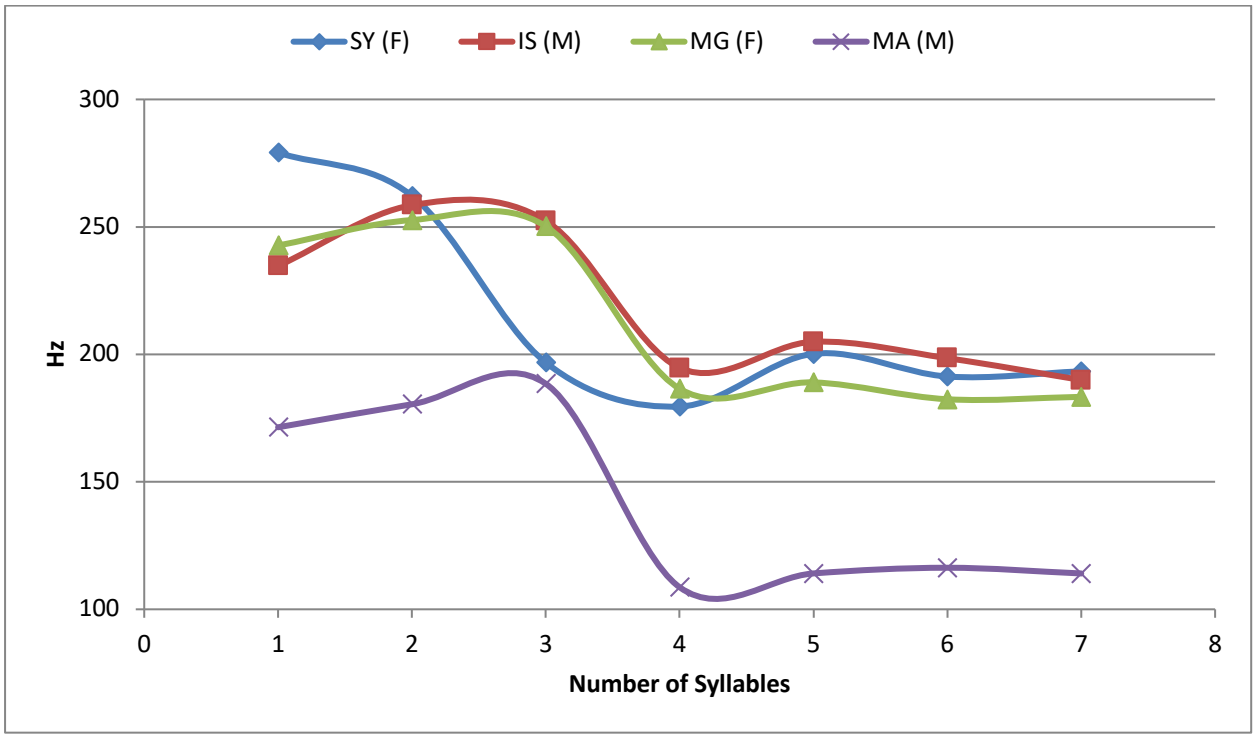


Figure 5.1(b): F0 values (in Hz) for sentence (62) showing points reached by every syllable per speaker

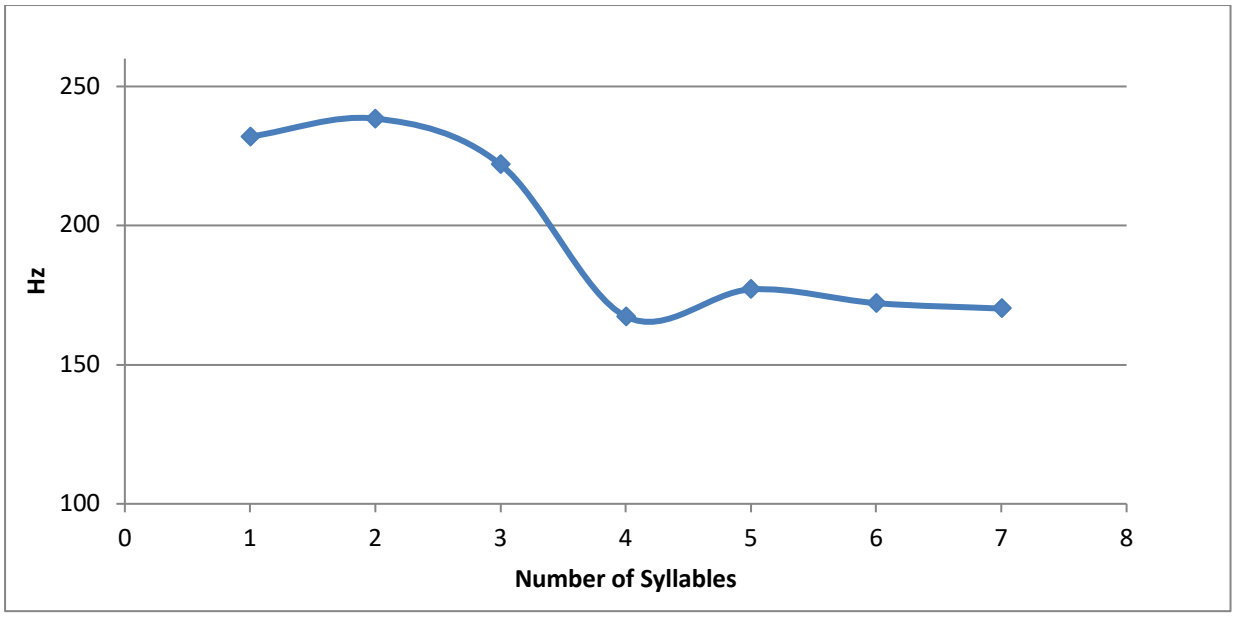


Figure 5.1(c): Mean F0 values (in Hz) for sentence (62) aggregated from 4 speakers

Figure 5.2 shows the pitch contour for sentence (63). The figure shows an alternation between Hs and Ls which creates a step-like contour. In this case, every intervening L-tone triggers downstep of the following H-tone except it is blocked by another tonal process. Tucker and Creider (1975) and Tucker (1994) refer to this form of downstep in Luo as explicit downstep.

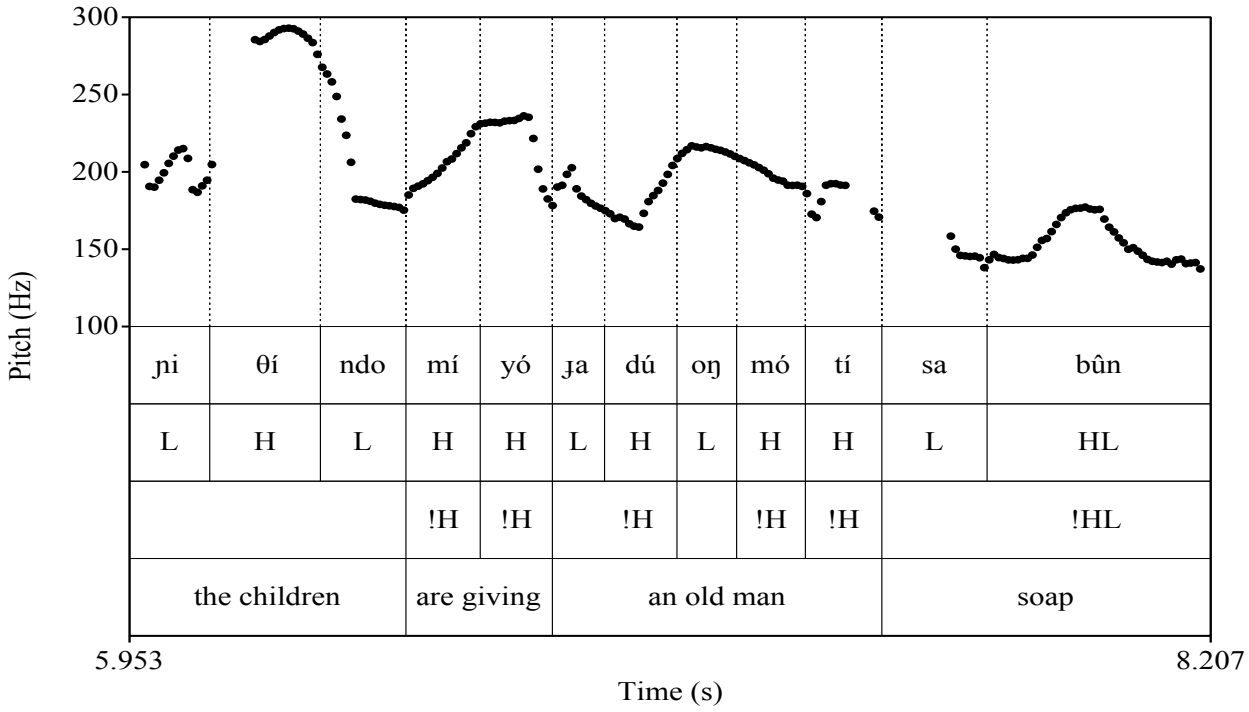


Figure 5.2 Pitch contour for a longer sentence (63) (with 12 syllables) showing automatic downstep triggered by intervening L tones; produced by a female speaker SY

Automatic downstep is a common trend in tonal languages especially those considered as terracing languages (see Clements 1979 and Huang 1985). In non-tonal languages like English downstep is also reported. However, Pierrehumbert (1980) accounts for downstep as an interpolation between H-tone and the targeted L-tone, a process which she considers to be phonetic rather than phonological. For her there is no underlying tonal contrast, but such

applicability is controlled by the quantitative relations (Pierrehumbert 1980:17). Such quantitative relations are well-explained by Liberman and Pierrehumbert (1984) who argue that automatic downstep (alias downdrift) is a function of exponential decay whereby pitch value decreases at the rate proportional to the current value towards a non-zero asymptote. In Autosegmental-metrical theory (Ladd 1996, 2008) downstep is viewed as a localized but iterated process. It can either be a gradual fall in pitch or a step-wise fall in subsequent Hs (Ladd 1996, 2008). In this thesis, I consider the gradual fall in pitch as declination and step-wise fall as downstep, which are respectively motivated by phonetic and phonological factors.

5.1.1.2 Non-automatic downstep

Besides automatic downstep, Luo also has incidences of non-automatic downstep. In this case, a sentence surfaces with Hs that are lowered with respect to the previous ones without an obvious L trigger. Consider sentence (64) below with its F0 traces in Figure 5.3.

64. L L H H !H !H

miyo ríémbó ndígá

woman PROG.ride bicycle

'A woman is riding a bicycle.'

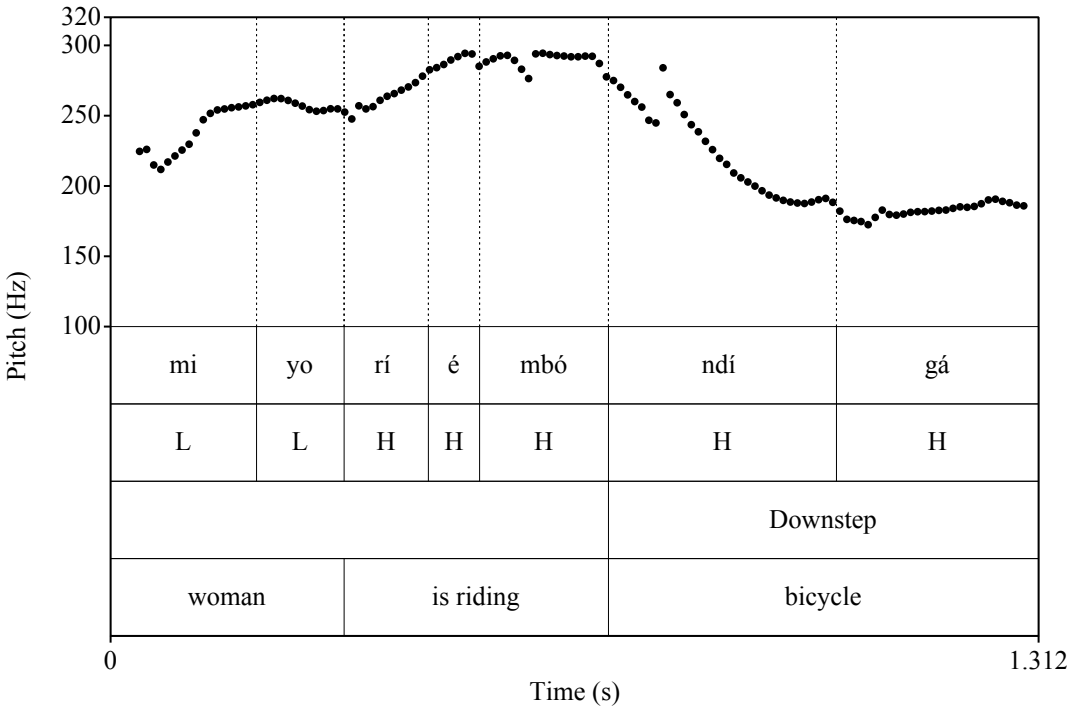


Figure 5.3 F0 curve for an all-High tone sentence (64) with non-automatic downstep at the NP left edge, produced by speaker SY

In Figure 5.3, there is a fairly dramatic fall in pitch affecting the two final Hs. Theoretically, it is assumed that there is a floating L from a deleted segment (or which can be historically derived) which triggers downstep (Yip, 2002; Gussenhoven 2004). If that is the case, it means the underlying structure for sentence (64) is L L H H H (L) H H (with a floating L in brackets) which then surfaces as L L H H H !H !H. This means the phonological rule applies before the phonetic interpretive rule (Clements, 1979). The latter is what we see on the surface as a physical representation of pitch. This phenomenon has been widely reported in tone terracing languages such as Akan (Clements, 1979; Genzel, 2013; Kügler, 2016).

However, it has been argued across languages that not all cases of non-automatic dowstep are triggered by a floating L (Clements, 1979). Lack of evidence for a floating L as a trigger of downstep is also reported in Bemba where downstep seems to adhere to the Obligatory Contour Principle (OCP) which prohibits adjacent lexical Highs (Kula & Bickmore, 2013; Kula & Hamman, 2016). In Tucker and Creider's (1975) early analysis of Luo intonation, they refer to this kind of downstep as implicit downstep. Their analysis is purely descriptive and hence does not make reference to any theoretical underpinnings to show any phonological rules involved. Regarding the case in (64), there is also no clear evidence of L-tone deletion or assimilation which would provide the phonological environment for the occurrence of a non-automatic downstep. In this case, it can then be argued that downstep at the left edge of an NP object *ndígá* 'bicycle'is triggered by boundary L% and not necessarily a floating L.

5.1.2 Declination

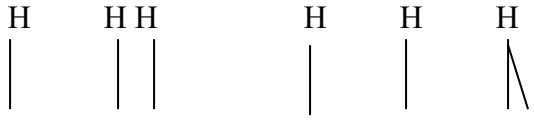
Another form of downtrend is declination. This is a gradual lowering of pitch in the course of an utterance. It involves narrowing of the pitch range which determines how tonal specifications are mapped into F0 values (Pierrehumbert, 1980; Gussenhoven, 2004; Downing & Rialland, 2016). As opposed to downstep, declination is a kind of lowering that is attributed to phonetic rather than phonological effect. It is said to have a universal effect affecting the whole phrase or utterance. This gradual fall in F0 is clearly seen in declarative sentences with sequences of like tones rather than mixed tones (Connell, 2001).

To establish whether Luo has declination or not, sentences with like or nearly-like tones were tested. It should be noted regarding verbs that there are always melodic tones inserted depending on the TAM. Thus, the grammatical pattern for verbs in their perfective aspect will be L HL L with a rise on the second syllable, and that of progressive aspect will be H H H or H H HL sentence finally. This patterning shows that there will always be an interfering H-tone even in a sequence of lows. This is the reason for having nearly-like tones and not totally identical. The following sub-sections explore the effect of declination in H and L sequences.

5.1.2.1 High tone sequence

In the earlier discussion we have seen that H tone sequence can be downstepped as a result of L% boundary. Apart from this effect, a sentence with all-High can also fall gradually indicating a phonetic effect, hence declination. Consider the sentence in (65) which contains all-High tones.

65.



món	tédó	réc	gí	yíén
women	PROG.cook	fish	with	firewood

'The women are cooking fish with firefood.'

Given that declination is a phonetic effect involving gradual fall in pitch, it is worth doing a quantitative test to evaluate the effect on the sentence. Hence, normalized pitch traces of the six syllables were obtained from four speakers. The results are presented in Table 5.1 below. In this table, the pitch value for each peak is indicated. The table also shows the difference in

pitch (Hz) between H1 and H2, H2 and H3…H5 and H6. The amount of pitch drop between consecutive Hs from H1 to H6 for every speaker is 18.1Hz for SY, 12.91Hz for IS, 16.57Hz for MG and 13.44Hz for MA. This makes an average difference of 15.25Hz in pitch drop aggregated over the four speakers.

	SY	△	IS	△	MG	△	MA	△
H1	286.2	30.26	245.55	-3.36	240.73	13.19	183.54	8.17
H2	255.94	28.19	248.91	8.83	227.54	27.37	175.37	16.82
H3	227.75	18.26	240.08	1.56	200.17	10.63	158.55	5.41
H4	209.49	13.17	238.52	50.61	189.54	27.91	153.14	9.1
H5	196.32	0.62	187.91	6.94	161.63	3.79	144.04	27.72
H6	195.7		180.97		157.84		116.32	
Mean △		18.1		12.91		16.57		13.44

Table 5.1 F0 values for H-tones in a six syllables sentence (65) and their differences in pitch drop (△) between H1-H2, H2-H3, H3-H4…, as produced by four speakers

Figure 5.4 presents the scaling of the averaged F0 values aggregated over the four speakers. This figure clearly shows the gradual fall in pitch from H1 to H6 which on average ranges between 250Hz and 150Hz. Following Gussenhoven's (2004) production code, what is important in declination is not the shape of the slope but the variation in utterance edges. In this case, the difference between the initial H and the final H is 100Hz.

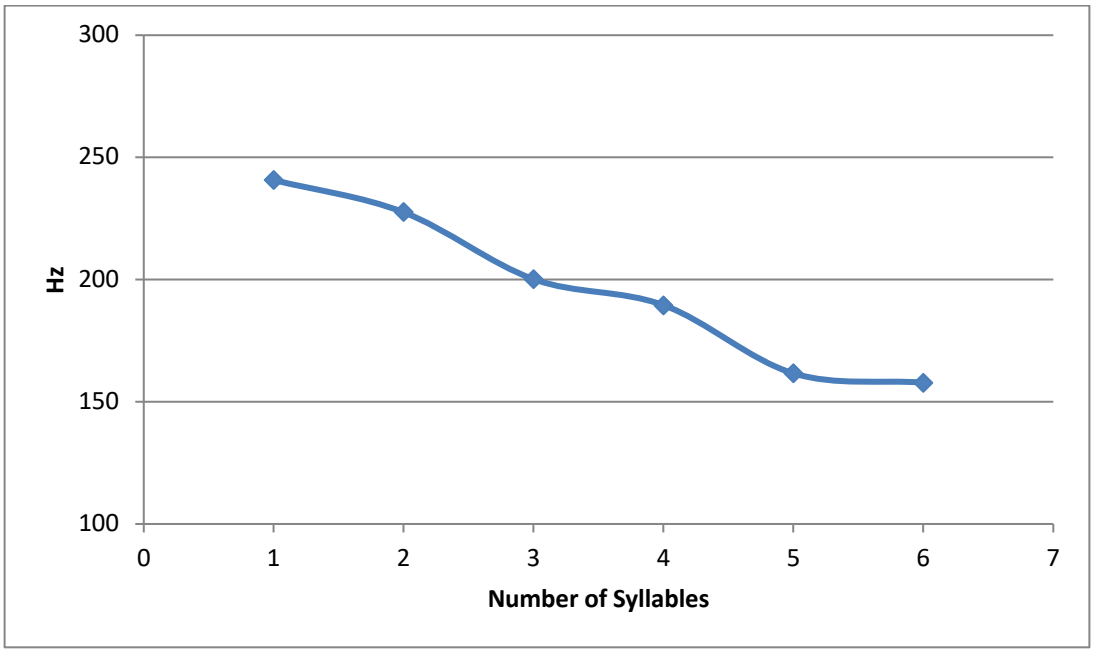


Figure 5.4 Average F0 values for sentence (65) aggregated from 4 speakers

5.1.2.2 Low tone sequence

A sequence of L-tones also portrays a declination effect. Consider sentence (66) below.

66.

L L	L HL L	L L L	L L
dayo	o-mîyo	wuoro	ŋuka
grandmother	3SG-PRF.give	father	porridge

'Grandmother has given father porridge.'

In order to measure the declination effect in a L-tone sequence, pitch values for each syllable from L1 to L10 were obtained from the same four speakers. Results are shown in Table 5.2 below. Pitch drop from L1 to L10 also show a declining effect with an average difference (△) of 3.58Hz. The amount of pitch drop for each speaker is 6.18Hz for SY, 2.8Hz for MG, 2.73Hz for MA and 2.62Hz for IS. SY has a greater amount of pitch drop with a significantly

steeper slope. As noted earlier, underlyingly the L-tone sequence is interrupted by a grammatical H. However, in this case, the H is neutralized into L or has a very slight rise because it is surrounded by L-tones (see the difference △ between T3 and T4 in Table 5.2).

	SY	△	MG	△	MA	△	IS	
L1	178.34	2.13	174.24	2.84	115.68	5.58	192.65	-2.26
L2	176.21	6.8	171.4	-2.04	110.1	-0.7	194.91	-1.26
L3	169.41	0.03	173.44	2.68	110.8	-0.32	196.17	3.64
HL4	169.38	1.95	170.76	5.78	111.12	6.57	192.53	2.91
L5	167.43	2.62	164.98	-1.05	104.55	-2.47	189.62	2.96
L6	164.81	6.71	166.03	-0.64	107.02	-0.84	186.66	0.62
L7	158.1	8.49	166.67	7.67	107.86	3.68	186.04	8.65
L8	149.61	13.38	159	4.68	104.18	10.7	177.39	6.29
L9	136.23	13.56	154.32	5.34	93.48	2.38	171.1	2.1
L10	122.67		148.98		91.1		169	
Mean △		6.18		2.8		2.73		2.62

Table 5.2. F0 values for L-tones in a ten-syllable sentence (66) and their differences in pitch drop (△) between L1and L2, L2 and L3, L3 and L4…, as produced by four speakers

Figure 5.4 presents the scaling of the averaged F0 values aggregated over four speakers, as shown in Table 5.2. In this figure, it is also clearly shown that there is a gradual fall in pitch from L1 to L10, whereby the first L is produced at around 165Hz while the final L is produced at around 132Hz, with the difference of 33Hz between the two edges. Generally, there is an overall drop in pitch from L1 throughtout L10.

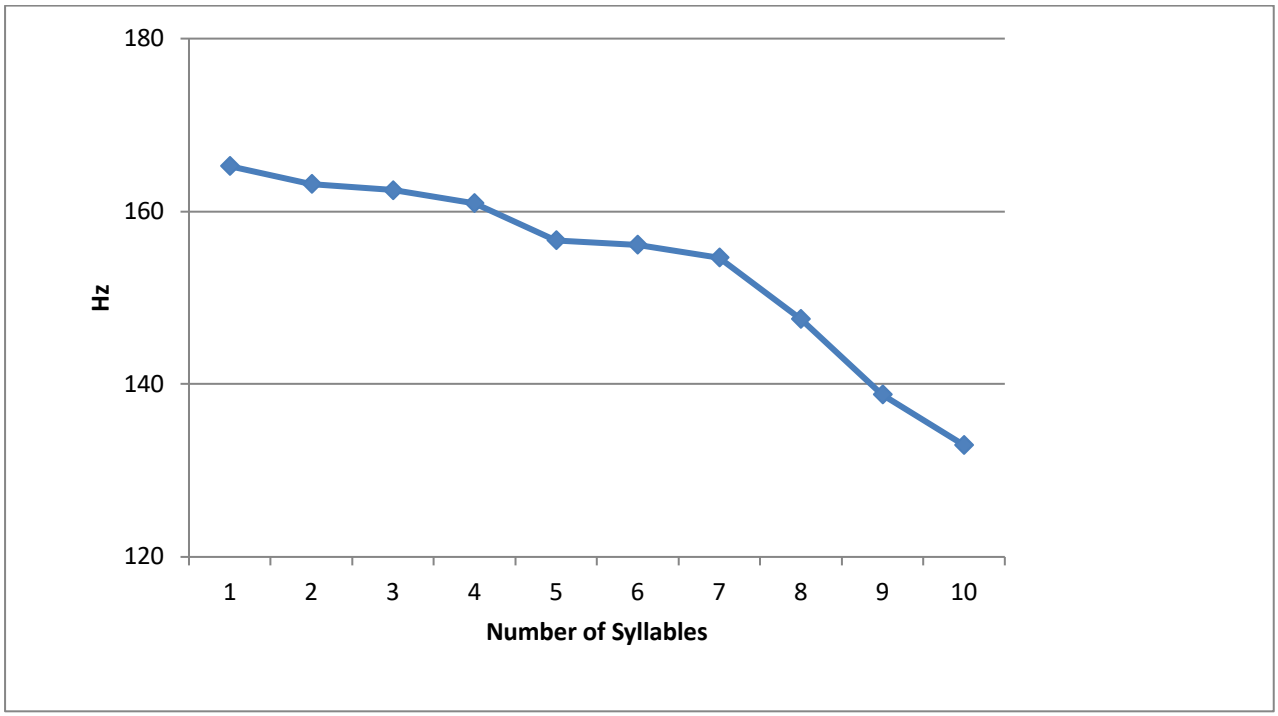


Figure 5.5 Average F0 values of sentence (66) aggregated from 4 speakers

The illustrations above show that both H-tone sequences and L-tone sequences in Luo exhibit declination. Tucker (1994) refers to this form of downtrend as downdrift which he also reports to occur in high tone sequences as well as in low tone sequences.

5.1.3 Final Lowering

Final lowering is another form of downward trend (distinct from declination) which affects the end of an utterance. Gussenhoven (2004) treats it as an extra dose of lowering. This dramatic effect at the right edge of a phrase is reported in intonation languages like English as well as in tone languages like Mandarin, Bemba and Chichewa. However, there are languages reported to lack final lowering or have it limited to certain tones. Lack of final

lowering is reported in Basaa (Makasso et al. 2016) and Konni (Cahill, 2016). Meanwhile Mambila (Connell, 2016) has final lowering on L-tones only.

In this section, I will examine whether Luo has final lowering or not. The analysis will follow the exponential decay model of Liberman and Pierrehumbert (1984). Although the model was designed to describe the relationship between a series of tonal items in English it can be applied to tone languages. The model stipulates that, with reference to non-final accents, final accents are produced lower than it would be predicted. It also argues that each number is a constant fraction of the previous one in a series of items, as formalized in (67).

67. $P_i+1 = (d \times (P_i - r)) + r$

This formalism is explained by Gussenhoven (2004:110) as follows: P is the peak value in Hz and *d* is a downstep factor between zero and 1. Gussenhoven (2004) recommends 0.7 as the mostly used *d* value. Every peak results from a multiplication of the F0 of the previous peak by *d*. This will cause every following peak to be lower than the preceding one. A constant *r* of 75Hz represents the reference line, to ensure that pitch does not go beyond the minimum. This formula can test the relationship between adjacent tones and as well as the relationship between final and non-final tones.

For the non-final tones, an idealized asymptote of an exponential decay specifically for final lowering is given in Figure 5.6 below. The image is based on Liberman and Pierrehumbert's (1984) model as represented in Truckenbrodt (2004). Rialland and Somé (2011) refer to the expected exponential decay as asymptotic with soft landing and the abrupt pitch lowering as hard landing.

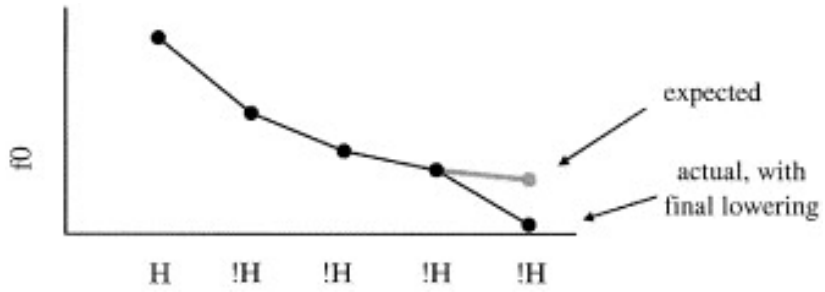


Figure 5.6 An asymptote showing expected F0 value and actual value, whereby the latter shows final lowering, after Liberman and Pierrehumbert (1984)

In order to investigate any final lowering effect in Luo, sentence (65) with Hs in non-final and final position was measured based on the formula in (67). Therefore if $r = 75$ and $d = 0.7$ and P1 = 240.73 then P2 = (0.7*240.73 – 75)) + 75) = 191.01. P3=175.38, P4=166.25, P5=162.71, P6=153.41. These values when scaled on F0 they produce an asymptote line, suggesting a lack of final lowering as shown in Figure 5.7.

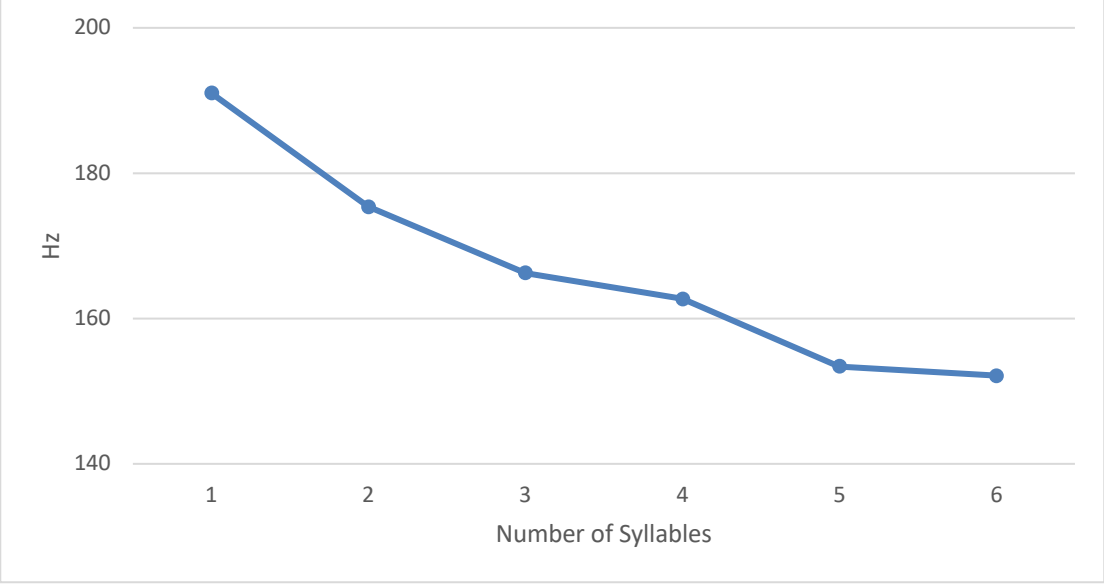


Figure 5.7 Averaged F0 H peaks (for sentence 65) in final and non-final position

Again, a similar measurement for exponential decay effect was conducted for sentence (66), which comprises L-tone sequence in final and non-final position. The results of the F0 scaling are illustrated in Figure 5.8. This figure shows a dramatic fall in pitch in the last three syllables. It is observed that the average difference of exponential decay of L-tones in non-final position is 0.56Hz. Meanwhile, at the final position, the average difference is 2.41Hz. This kind of abrupt lowering, whether it affects a single TBU or a range of TBUs, can be attributed to final lowering. This contrasts with declination which is more gradual as a phonetic effect, as indicated in the first six syllables.

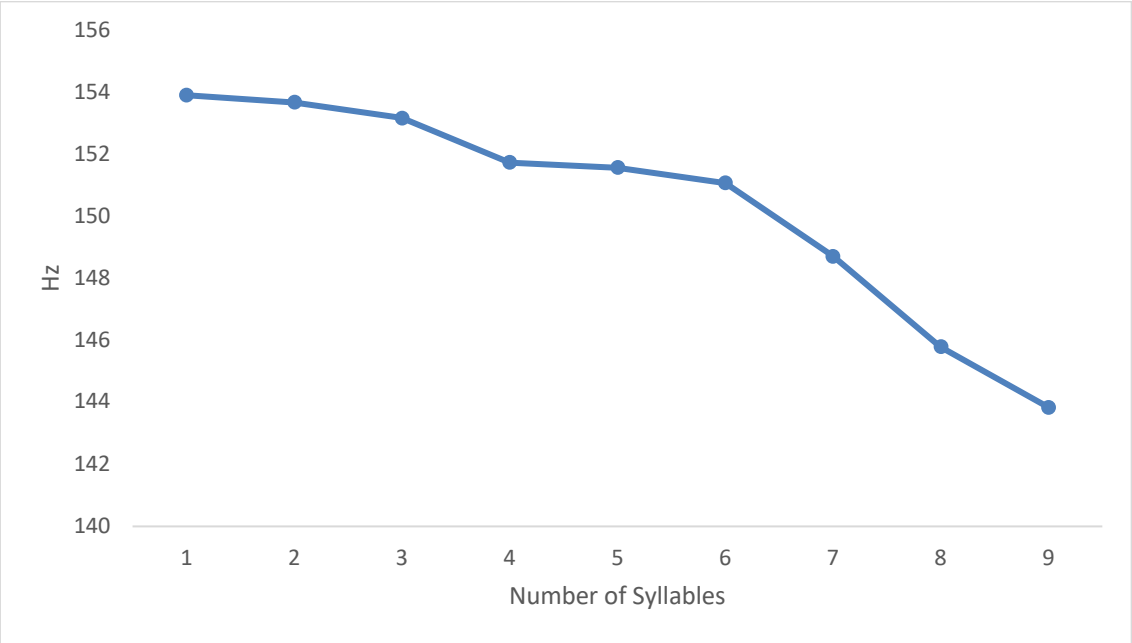


Figure 5.8 Average pitch values of L-tones (for sentence 66) in non-final and final position

Thus, final lowering in Luo is limited to L-tones only. A similar case is reported in Mambila by Connell (2016). In the case of H-tones, there is normal downstep effect at the final position as it would be in any downstep case, as presented in Figures 5.1 and 5.2. Note that in Figure 5.2 the final lexical tone which is HL also undergoes downstep which can be due to

the word being in the final position or due to the preceding L. This means even contour tones, so long as they have H tones, will be downstepped sentence finally. According to Gussenhoven (2004) final lowering can be grammaticalized distinctly from downstep but it can also be tied up with downstep. He gives an example of Dinka (a Nilotic language) which always downstep utterance final H-tone syllables, regardless of the preceding tone, to mark finality. Thus, in Luo, downstep can occur as a discrete process or can be tied up with a L% boundary utterance finally, not necessarily final lowering.

To sum up, declaratives generally have a falling contour realized by register shift as in other tone terracing languages. This can be in form of downstep or declination whereby the former is phonologically motivated and the latter is phonetically motivated. There are also final effects observed, being associated with boundary L%. In this, final lowering, as an abrupt final fall is mostly observed in utterances which end in L-tone. Meanwhile H-tones utterance-finally are downstepped, exhibiting a flattened F0.

5.2 Questions

The second part of this chapter investigates the intonation of questions. Two main types of questions will be investigated. These are polar questions, also known as yes-no questions, and content questions. A full list of the questions used in this investigation is found in Chapter 4 section 4.3.1. Generally, the syntactic and phonological compositions of the questions used resemble that of the statements as presented in section 5.1 above.

5.2.1 Polar questions

Structurally, Luo polar questions use the same structure as a declarative utterance with an optional question particle. There are several particles which can be used. These include *koso, donge, kare, bende*, and *ni* but the commonly used one is *be* a short form and a variant of *bende.* In this work, I will restrict my investigation to *be.* It should be noted that *be,* as well as other particles, can occupy different positions in a sentence. For example, in (68) and (69) the particle *be* is at the initial position functioning as a question particle while in (70) *be* is not in initial position and here functions as a question particle or as a conjunction which means *also* or *too*. The difference in meaning will depend on the chosen intonation. As already presented above, the hypothesis is that a statement interpretation will have a low pitch range and a question interpretation will have a high pitch range. This will be tested shortly.

68. be món tédó réc gí yíén
    Q women PROG.cook fish with firewood
    'Are the women cooking fish with firewood?'

69. be dayo míyó wuoro ɲuka
    Q grandmother PROG.give father porridge
    'Is grandmother giving father porridge?'

70. má be ɟadúòŋ adìér
    this Q/CONJ old man really
    'Is this an old man really?'/'This is also an old man really.'

In the work done by Rialland (2007) on yes-no questions in African languages it was shown that a large number of languages have lax question prosody, which is characterised by falling pitch contour. Lax prosody may occur in isolation or in combination with other phonetic features. In Luo it is observed that polar questions have pitch register expansion (PRE), as observed in Figure 5.9 and 5.10. In these figures, PRE is potrayed with an increase in F0. Phonologically, this is represented by an initial -H at the left edge. By superimposing a polar question over a declarative counterpart, we can clearly see which parts are raised and which ones are not. This representation is borrowed from Kula and Hamman (2016). PRE can affect the entire or some parts of the clause. In Figure 5.9, the part of the clause which constitutes the subject and the verb has expanded F0 while the other part which constitutes the object and the preposition phrase is relatively constant in relation to the statement. This shows that polar questions and statements can have a boundary L%, which is manifested as final lowering when the final tone is low or a downstepped !H when the final tone is high. In Figure 5.9 the L% boundary targets a number of syllables.

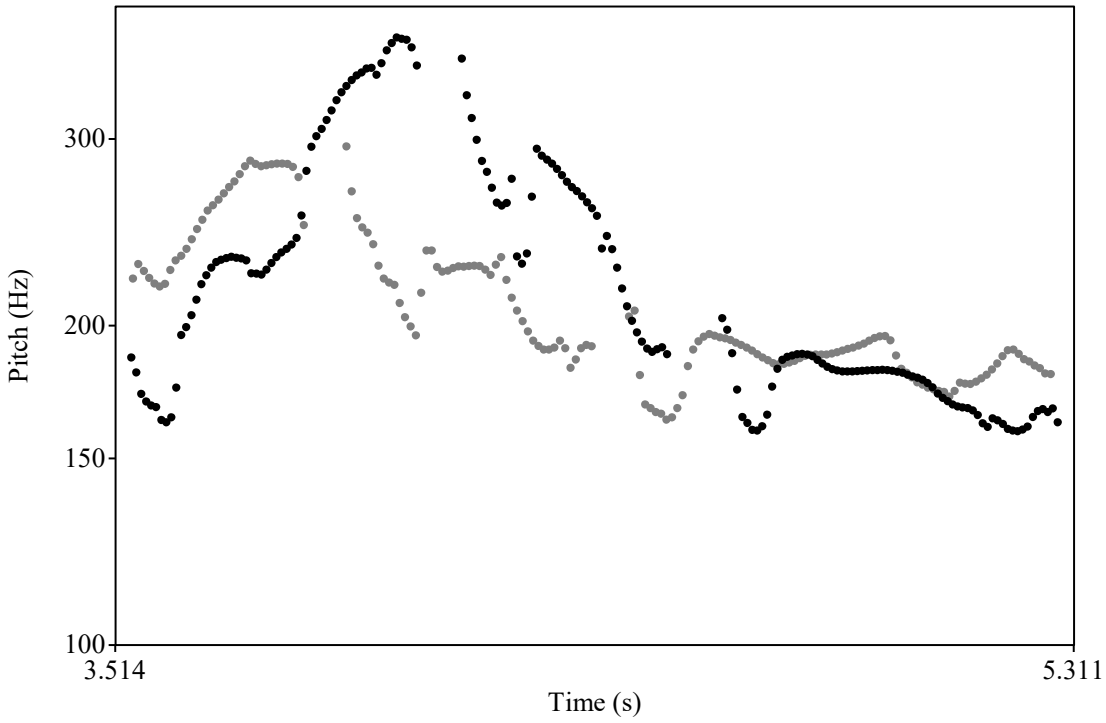


Figure 5.9 A comparison between a declarative sentence (grey) *món tédó réc gí yién* (example 65) and a yes-no question (black) *be món tédó réc gí yién* (example 68), produced by speaker SY

In Figure 5.10, the polar question also has a falling contour but with an expanded pitch throughout the clause. There is also a boundary L% targeting the final syllable. Based on these examples it can be argued that Luo polar questions use a combination of lax prosody characterised by a falling contour or low ends and pitch register expansion. According to Ohala (1983) as well as Gussenhoven (2004) the use of high pitch register to differentiate questions from statements is the result of grammaticalization of the informational interpretation of the frequency code.

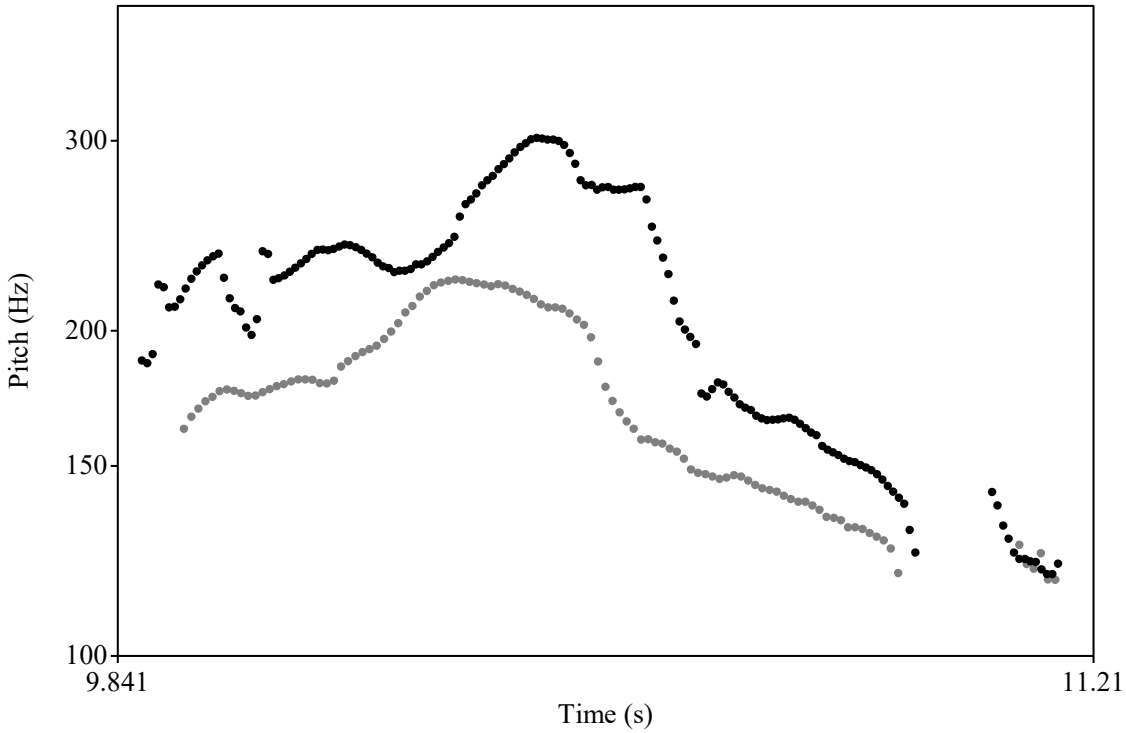


Figure 5.10 A comparison between a declarative sentence (grey) *dayo míyó wuoro ɲuka* and yes-no question (black) *be dayo míyó wuoro ɲuka* (example 69), produced by speaker SY

Apart from PRE, there is also the possibility of having a rise-fall (HL) contour sentence finally. Although this type of question prosody is not native to Luo it was produced by two speakers among the four. In spontaneous speech this melody was not noticed. It only appeared in read text. Therefore, the assumption is that perhaps speakers sometimes tend to mimic the Swahili question prosody by producing a HL melody.[13] This assumption is based on the fact that Swahili is a dominant second language in the area resulting in some language contact effects, which are not limited to question intonation but also to other linguistic aspects such as syntax.

[13] This is based on my own observation and knowledge of the Swahili language.

Figure 5.11 illustrates the F0 traces of a yes-no question with a HL final, against its declarative counterpart. Here we see that the pitch for a yes-no question is relatively similar to that of a statement, with the only contrast being that a HL melody is added to the final TBU. Although this structure was produced by a male speaker, MA, it cannot be treated as possibly gender (male) related-variation because the other male speaker, IS, produced the same questions using PRE as the female speakers did. Another point to note is that speaker MA somehow consistently produced HL melody with imperfective aspect only and rarely with perfective aspect. So, based on these data I consider the use of HL melody utterance finally as an alternative way of asking yes-no questions in the language.

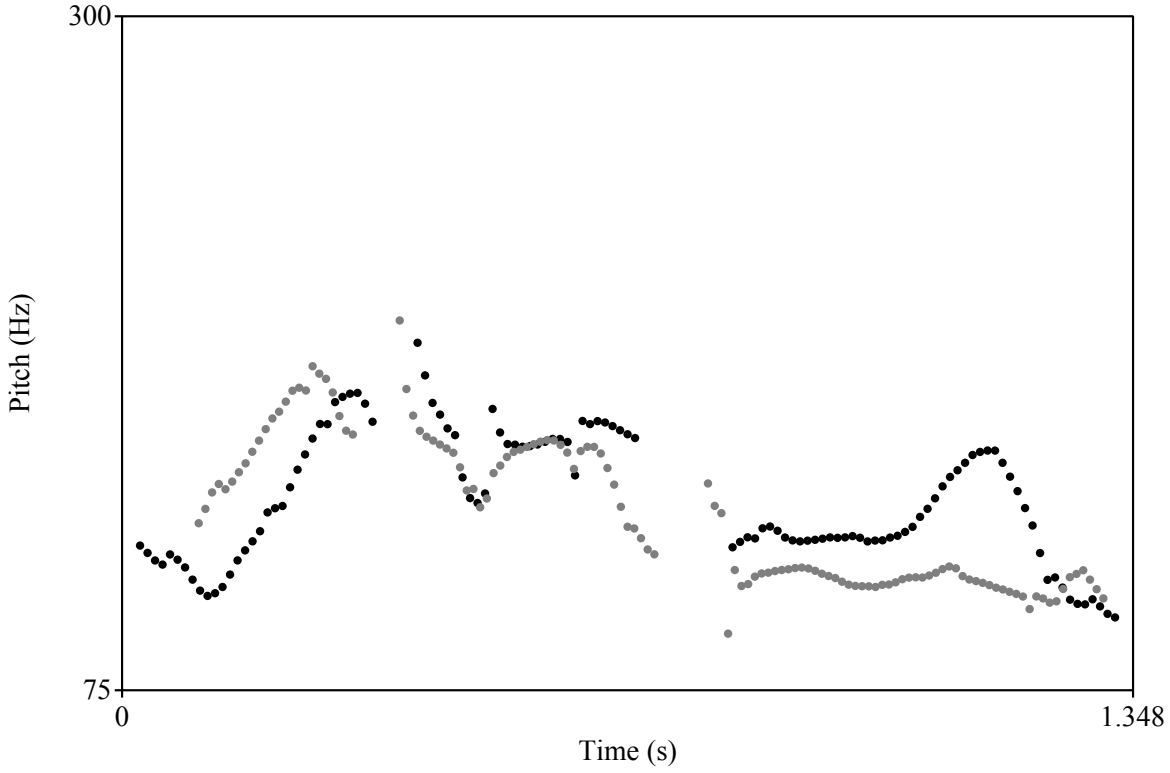


Figure 5.11 A comparison between a declarative sentence (grey) *món tédó réc gí yién* (example 65) and a yes-no question (black) *be món tédó réc gí yién* (example 68) showing final HL melody, produced by speaker MA

Thus, given the illustrations above, at least two intonational structures for polar questions are identified. These are PRE and a HL%. Both intonation strategies are not uncommon in African languages. For example, PRE is reported in Bemba (Kula & Hamann, 2016) and in Chichewa and Tumbuka (Downing, 2016) and a combination of PRE and HL melody in Embosi (Rialland & Aborobongui, 2016). Additionally, in Bemba, Chichewa and Tumbuka, downstep is said to be totally suspended in yes-no questions. This is however different in Luo whereby it is observed that downstep of H is maintained even when pitch register expands. Compare Figure 5.1(a) and Figure 5.12 below. The former represents a contour for a declarative sentence while the latter is a contour for a yes-no question, but both exhibit downstepped Hs.

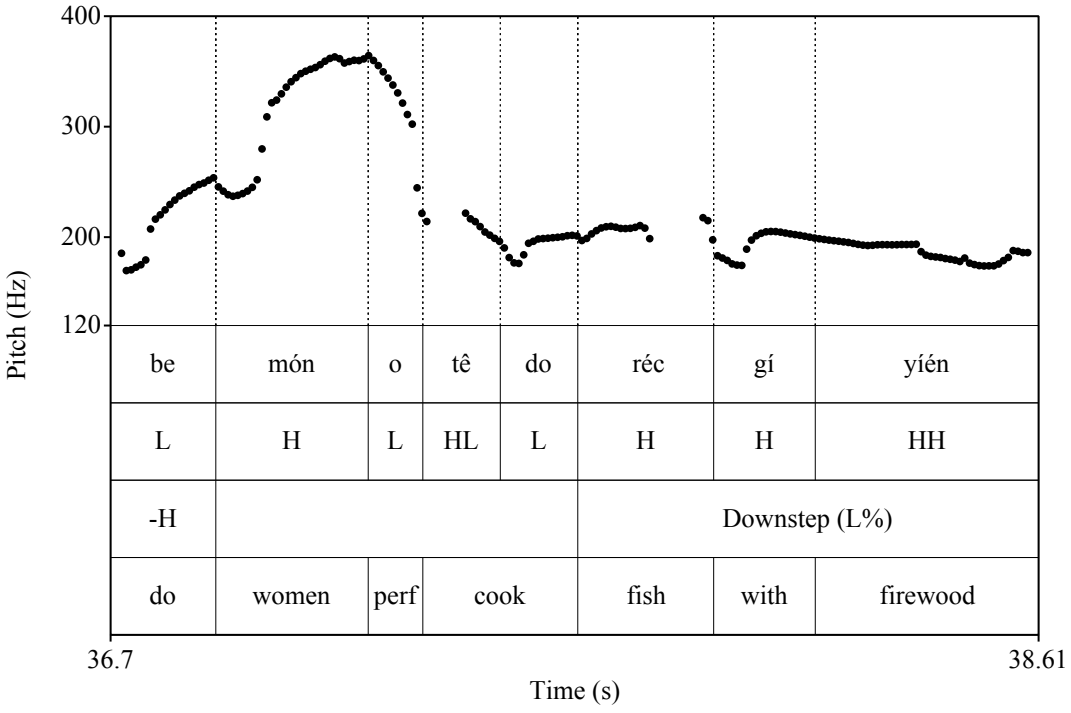


Figure 5.12 F0 contour of a yes-no question showing PRE with downstepped Hs utterance finally, produced by speaker SY

5.2.2 Content questions

Luo content questions, as in most African languages, use in-situ question words to express interrogativity. In Luo, content questions use different forms equivalent to wh-words in English. These words are *ŋa* which means 'who', which can be used for both subject and object questions. However, when *ŋa* 'who' is used as a subject question, a particle *má* will be added to create a cleft construction, *ŋa má,* which literally means 'who is it'. This is because subjects are never questioned in-situ as other constituents unless they are clefted. Other forms are *aŋó* for 'what', *máné* for 'which', *karaŋó for* 'when', and *káŋé* for 'where'. In this discussion we will consider the subject and object questions: *ŋa* (as shown in examples 71 and 72) and *aŋó* (as in 73); as well as adverbial questions using *káŋé* and *karaŋó* (as in 74 and 75, respectively) **.** The respective F0 contours for these questions are presented in Figures 5.13 to 5.17. It should be noted that the same question can be used for eliciting focus. The discussion about focus is presented in chapter 6.

71. **ŋa** má ŋíéwó obámblá
who REL PROG.buy dried fish
'Who is buying dried fish?'

72. dayo míyó **ŋa** ɲuka
grandmother PROG.give who porridge
'Who is grandmother giving porridge?'

73. ɟapur bíró gi **aŋó**
farmer PROG.come with what
'Whatis the farmer coming with?'

74. ɟapúóɲɟ ríémbó ndígá **káɲé**
teacher PROG.ride bicycle where
'Where is the teacher riding a bicycle?'

75. ɟapúóɲɟ lémó **karaŋó**
teacher PROG.pray when
'When is the teacher praying?'

Like polar questions, content questions in figures 5.13 to 5.17 exhibit pitch range expansion at the left edge, which is marked by -H boundary. This is evidenced by the L-tone(s) sentence initially produced at or above 200Hz. Normally, as seen in the previous sections, L-tones are produced below 200Hz. Furthermore, in these figures it is observed that the question word is prosodically non-prominent as it is portrayed differently depending on its position in a sentence. For example, in Figure 5.14, the low-toned question word *ŋa* 'who' is in non-final position and is affected by declination. The question word *aŋó* 'what' in Figure 5.15 is affected by downstep because it is in the final position of the clause. The same applies to *káɲé* and *karaŋó* which are also dowsntepped in the final position. As already discussed earlier in declaratives, L-tones are affected by final lowering while H-tones are downstepped

utterance-finally. The same is also observed in questions, as a result of final effects triggered by boundary L%.

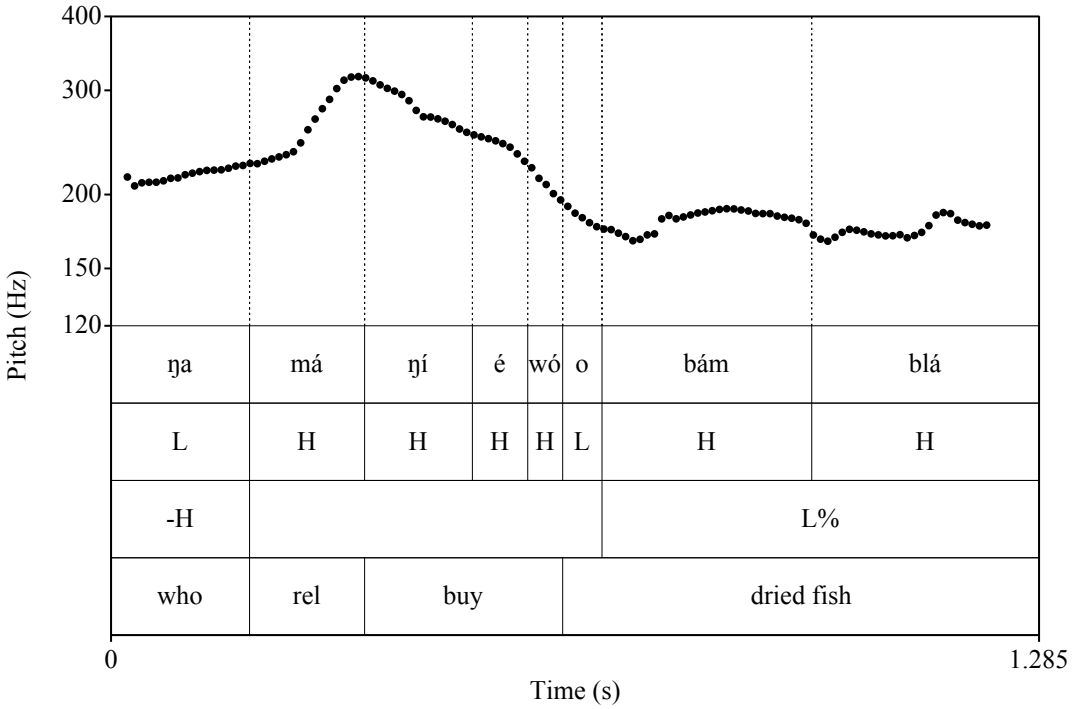


Figure 5.13 PRE in a content question (example 71) with a low-toned question word at the left edge exhibiting -H boundary, produced by speaker SY

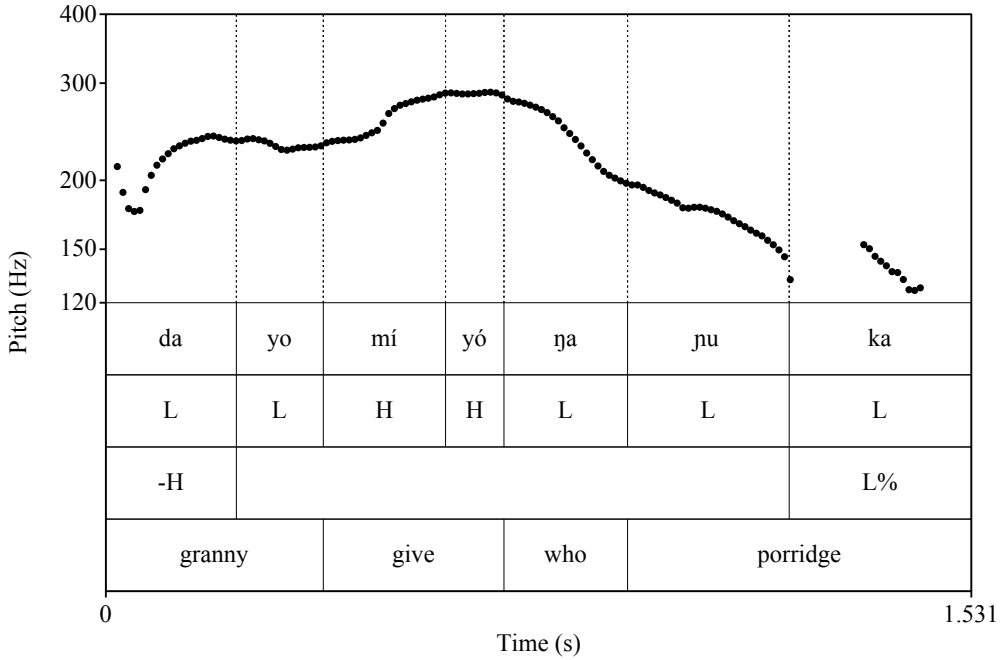


Figure 5.14 PRE in a content question (example 72) with the question word in non-final position affected by declination, produced by speaker SY

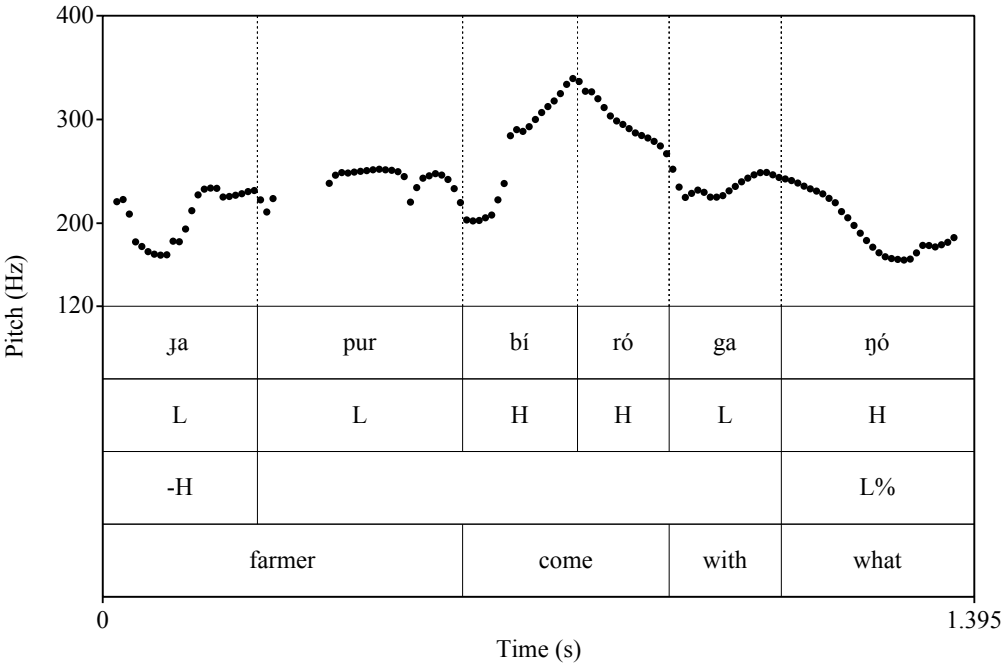


Figure 5.15 PRE in a content question (example 73) with the question word at the final position exhibiting a downstepped H, produced by speaker SY

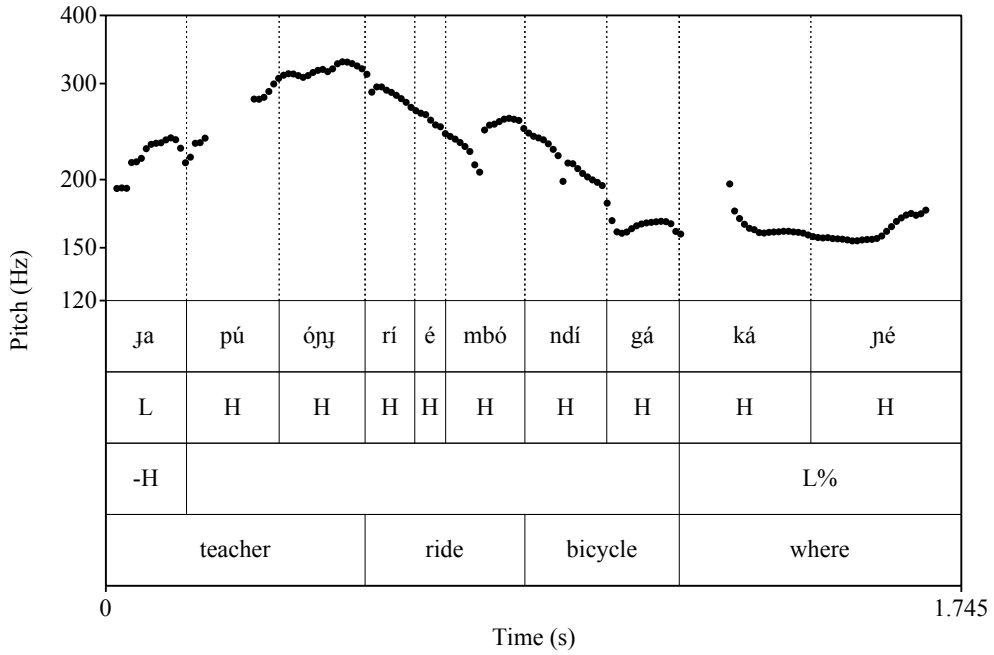


Figure 5.16 PRE in a content question (example 74) with the question word in final position affected by downstep, produced by speaker SY

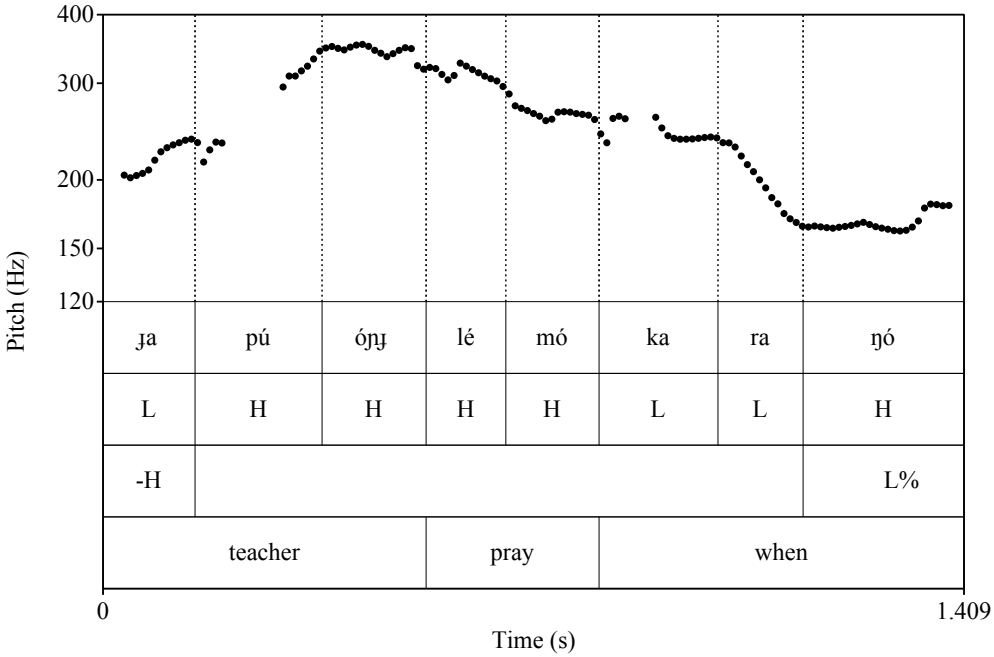


Figure 5.17 PRE in a content question (example 75) with the final syllable of the question word affected by downstep, produced by speaker SY

Although PRE is the main prosodic feature for content questions in Luo, there can also be a HL% melody. This affects the final syllable in an utterance as portrayed in Figure 5.18. In this case PRE expansion is also applied but downstepping of H-tones is suspended. A HL% melody is reported in some African languages. For example, in Fipa, Riedel and Patin (2011) report that wh-questions have a H(L%) melody in the final syllable.

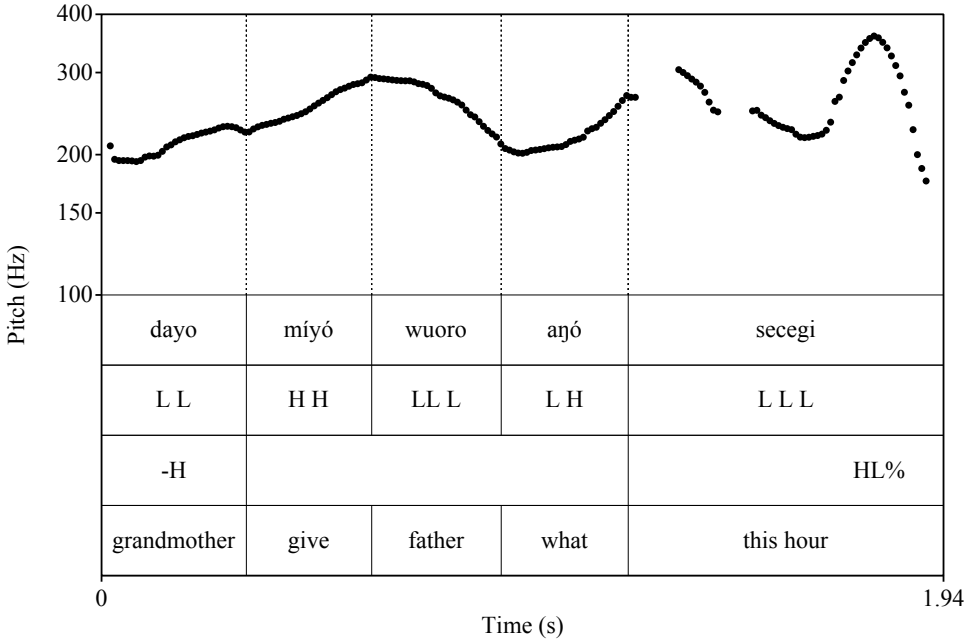


Figure 5.18 Content question with HL% utterance finally, produced by speaker IS

As it was for polar questions, content questions in Luo also have pitch range expansion (PRE), suggesting that prosodically polar and content questions have the same features. HL% is also seen as an alternative prosody for content question.

5.3 Conclusion

In this chapter a detailed discussion of the intonation of declaratives and questions is presented. It is noted that declaratives as well as questions exhibit downtrends. These are

automatic downstep, non-automatic downstep, declination and final lowering. In this study, I have considered automatic downstep as a local phonological effect resulting from an interaction between adjacent tones particularly H and L. Meanwhile, non-automatic downstep in the cases presented portrays an interaction between lexical tones and intonation boundaries. Tucker and Creider (1975) and Tucker (1994) present these two forms of downstep as explicit and implicit downstep, respectively. Declination, on the other hand, is portrayed as a phonetic effect with gradual fall in pitch affecting an entire phrase or utterance and this is noted in all-High and all-Low sentences. Tucker and Creider analyse this kind of gradual fall in like-tone contexts as downdrift. In their analysis they argue that downdrift has a fall that is gentle, constant and continuous, which are the characteristics that correspond to declination. Furthermore, final lowering as a dramatic fall in pitch sentence-finally is limited to Low tones. H tones are downstepped in final position, which, according to Gussenhoven (2004) is a way of grammaticalizing final lowering. Generally, the right edge of both declaratives and questions is characterized by a L% boundary (in form of final lowering or downstepped H).

Regarding questions, both polar and content questions have expanded pitch register (PRE) phonologically marked by -H at the left edge of the clause. Despite the noted PRE, questions also have falling pitch whereby H tones are downstepped and final syllables are lowered, as in declaratives. Additionally, some speakers superimpose a rise-fall (HL) melody sentence finally in both polar and content questions, a phenomenon which is caused by language contact from the dominant second language, Swahili. For the case of question words, there is no prosodic prominence applied to them.

Chapter Six

Intonation of Focus, Topics and Dislocations

6.0 Introduction

This chapter discusses the syntactic and prosodic aspects of focus constructions, topics and dislocations as part of information structure. Information structure alias information packaging is about how information is packaged in a sentence, reflecting the interlocutors' shared knowledge (Güldemann et al. 2015). There are different ways in which information structure is presented cross-linguistically. This ranges from morphological marking to syntactic manipulation of a sentence. The latter involves dislocations and clefts (Güldemann et al. 2015). In addition to this type of morpho-syntactic marking, there might be special prosody such as change in pitch or duration to make the focused constituent prominent. The general assumption, particulary in intonation languages such as English and Dutch, is that information structure is prosodically prominent (see Selkirk, 2002; Gussenhoven, 2004). Furthermore, in these languages, it is reported that presentational (non-contrastive) and contrastive focus have different prosodies (Selkirk, 2002; Gussenhoven, 2004; Hartmann, 2008). Gussenhoven (2004) adds that some languages use pitch register raising or delayed peaks. Pitch register raising is also reported in tone languages such as Mandarin. In Mandarin, narrow focus has expanded pitch register while the non-focused part is compressed. In broad focus neither expansion nor compression is applied (Peng et al., 2005). In recent studies of intonation in African languages, Downing and Rialland (2016) point out that a great number of languages do not have focus prosody. They point out further that some Bantu languages use prosodic phrasing to mark the presence of a focused or topicalized constituent. In such languages the same cues used to identify other forms of prosodic phrases

are also used to identify focused or topicalized constituents. An example of such case is present in Shingazidja (see Patin, 2016). Given these cross-linguistic variations it is worth investigating how information is structured in Luo. Attention will be given to both syntactic and prosodic aspects.

The chapter is organized as follows. Section 6.1 deals with discourse new information obtained from answers to wh-questions, while section 6.2 looks at contrastive focus. Section 6.3 concentrates on clefts and section 6.4 covers dislocations, investigating right dislocations, left dislocations and topicalized subjects. Finally, section 6.5 concludes the chapter.

6.1 Wh-question focus[14]

Wh-question focus is the most widespread form of focus, where an answer corresponding to a wh-word provides discourse-new information. Wh-question focus is also known as presentational focus (Selkirk, 2002) or non-contrastive focus (Hartmann, 2008). It can be narrow or broad. Broad focus applies when the whole expression provides new information, while narrow focus refers to when only a small part of a sentence provides new information (Gussenhoven, 2004). In the Luo cases presented below, the sentences used had narrow focus in medial and final positions. In Luo, as in many African languages, subjects cannot be focused sentence-initially unless they are clefted. Therefore, focus in this position will be discussed under clefts and topics in section 6.3 and sub-section 6.4.3, respectively.

[14] 'Wh' is used as a generic term to refer to question words in Luo where they take various different forms as will be seen in the following discussion.

As noted earlier question words in Luo are basically in-situ although they can be fronted. This case is also presented in section 6.3 under clefts. Being a so-called wh-in-situ language, the position of a focused constituent then directly correlates with the position of a wh-word. Let us consider the sentences in (76) and (77). Sentence (76a) is a wh-in-situ question which prompts object focus in (76b) and sentence (77a) is also a wh-in-situ question prompting adverbial focus in (77b). Note that focused constituents are in brackets in the Luo examples below and capitalized in the English gloss.

76. (a) dayo míyó ŋa ɲuka
grandmother PROG.give who porridge
'To who is grandmother giving porridge?'

(b) dayo míyó [wuoro]$_{Foc}$ ɲuka {Object Focus}
grandmother PROG.give father porridge
'Grandmother is giving FATHER porridge.'

77. (a) ɟapúòŋɟ lémó karaŋó
teacher PROG.pray when
'When is the teacher praying?'

(b) ɟapúòŋɟ lémó káwúònó
teacher PROG.pray today
'The teacher is praying TODAY.'

The focused word *WUORO* ‘father’ in (76b) has lexical L-tones while *KÁWÚÒNÓ* ‘today’ in (77b) comprises of HHLH pattern. Regarding prosodic realization, there is no prominence on the focused words. However, when a declarative sentence is superimposed onto the corresponding focused sentence, the results show that the latter has overall pitch raising. This is illustrated in Figures 6.1 and 6.2 below. Note that the declarative is in grey and the narrow presentational focus in black.

Figure 6.1 shows that the initial L-tones of a declarative (in grey) are produced at a lower pitch register, below 200Hz. In this case, only the H-tones on the verb *mîyó* which also spread to the first syllable of the focused constituent *wúoro* are produced above 200Hz. On the other hand, a narrow presentational focus sentence shows a similar pitch contour to that of a neutral declarative, except that it has overall pitch register raising with the initial low tones produced above 200Hz. In both cases, *wuoro* is affected by declination.

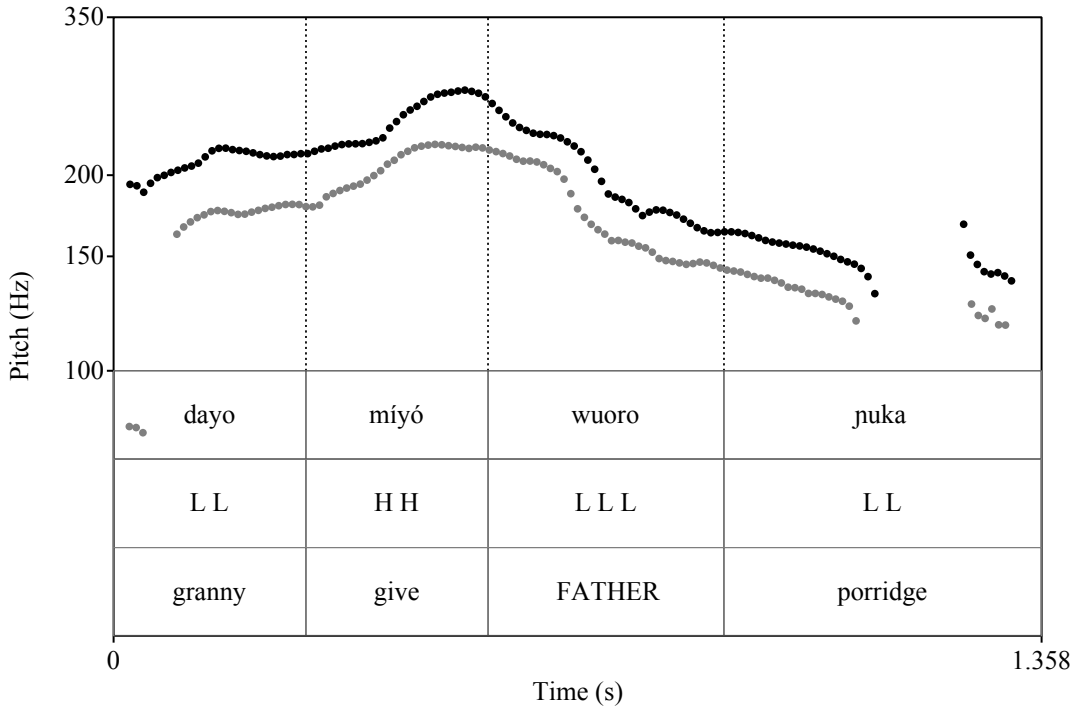


Figure 6.1 A superimposed image of a declarative sentence (in grey) and a focused sentence (76b) (in black) with object focus, showing differences in overall pitch register, produced by the same female speaker, SY

Similarly, in Figure 6.2 the sentence with adverbial focus has overall pitch raising, while the declarative is produced at a lower range. Low tones are generally produced below 200Hz while H-tones are produced above 200Hz. This specifically happens with the female speakers who have wider pitch range. Male speakers also exhibit similar pitch contours but at a lower range.

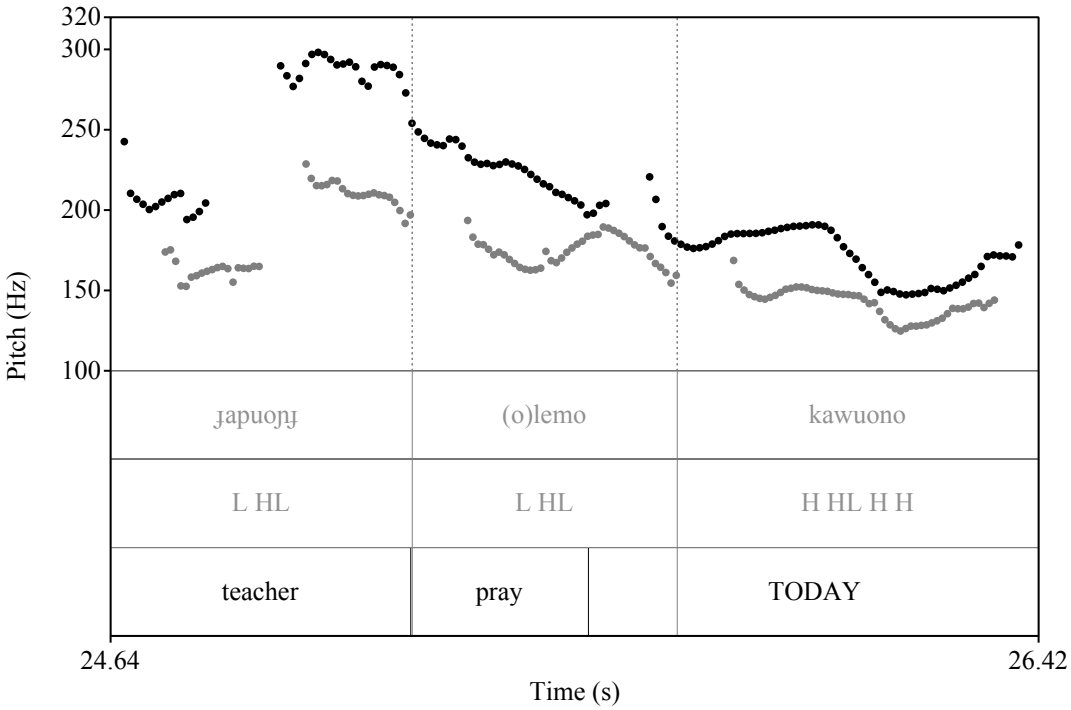


Figure 6.2 A superimposed image of the declarative (in grey) and adverbial focus (77b) (in black), showing differences in overall pitch, produced by the same female speaker, SY

With these examples, we can conclude that presentational focus has no discrete prominence marking. The raising of pitch at the left edge of a clause indicates the presence of a high boundary (-H) resulting in raising the pitch register of the entire clause, yielding a global effect as in the question intonation discussed in the previous chapter. This phenomenon is not uncommon in tone languages. In Mandarin, for example, pitch expansion is applied on the focused constituent while the non-focused constituent is compressed (Peng. et al., 2005). In Bemba pitch expansion occurs on the constituent preceding the focus which is usually a verb (Kula & Hamann, 2016).

As discussed previously in section 5.2, change in pitch range is also present in polar and content questions in Luo. It is said that question words are inherently focused, meaning there is an established correlation between question words and focused constituents (Kula & Hamann, 2016). Since question words have no prosodic prominence in Luo, it is therefore not surprising that focused constituents also lack prominence.

6.2 Contrastive focus

Apart from wh-questions, focus can also occur in controlled contexts such as coordinated parallel phrases, constructions where items can be selected from a set of alternatives and correction sequences (Hartmann, 2008). All these types of focus fall under the umbrella term contrastive focus. The first example of contrastive focus is given in (78). This sentence is composed of two clauses which are coordinated by *to* 'and'. In this case an exhaustive question-answer pair is used. In both clauses, focus is on the VP. Note that the VP in the first clause constitutes a verb and an object, and in the second clause it constitutes the verb only.

78. Qn: wuoro tímó aŋó to ɲaθí tímó aŋó
father PROG.do what CONJ child PROG.do what
'What is the father doing and what is the child doing?'

Ans: wuoro [dwáló kuon]$_{Foc}$ to ɲaθí [túgó]$_{Foc}$
father PROG.cook ugali CONJ child PROG.play
'The father IS COOKING UGALI and the child IS PLAYING.'

The pitch contour for the answer in (78) is illustrated in Figure 6.3. In this figure, we see the pre-focus constituents *wuoro* 'father' and '*to ɲaθî* 'and child' despite having L-tones are produced at a higher pitch, above 200Hz, suggesting pitch register raising. At the right edge of the first clause we notice a LH% boundary (with a slight rise which indicates continuation) and a downstepped H on the second clause, which also indicates a boundary L%. At the right edge of the second clause there is pitch rest indicated by a L-toned word resetting to 230Hz.

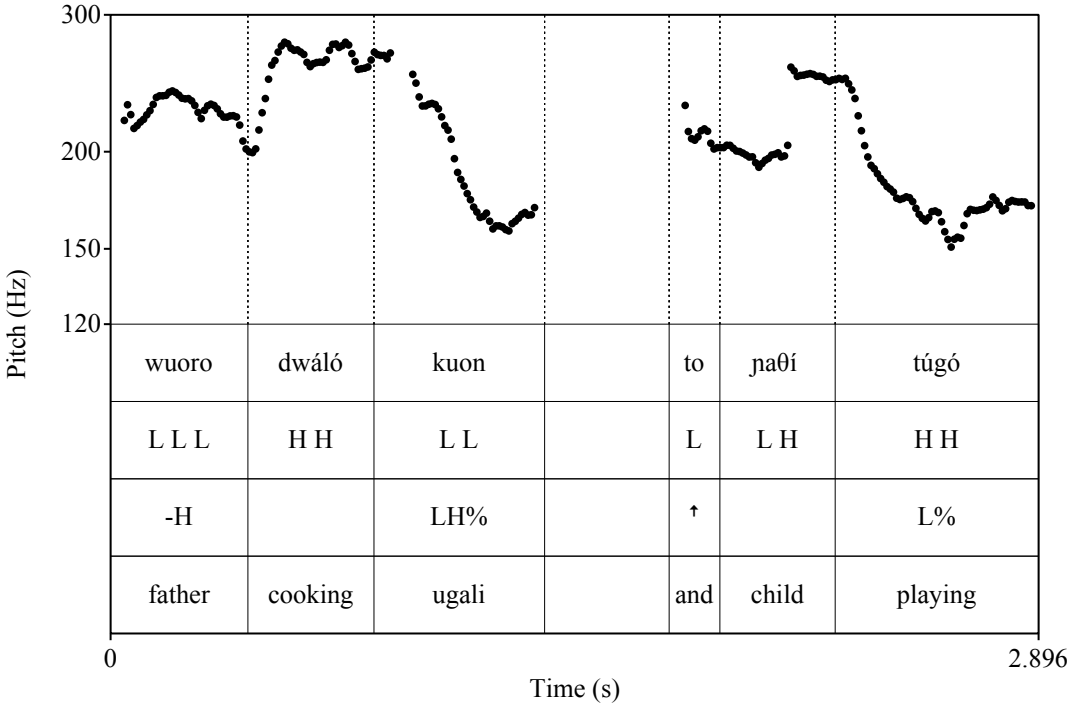


Figure 6.3 Pitch contour for the answer in (78) with pitch raising on the pre-focused word(s), produced by speaker SY

Another form of contrastive focus is alternative focus. In alternative focus a set of alternatives is given from which a relevant item is selected. This can also be achieved using an exhaustive question, as in (79). The pitch contours for the focus sentence in (79) is illustrated in Figure 6.4.

79. Qn: ó-tédó mcêle koso ó-tédó kuon
SM3SG-cook rice or SM3SG-cook ugali
'Is he cooking rice or is he cooking ugali?'

Ans: ó-tédó kuon
SM3SG-cook ugali
'He is cooking UGALI.'

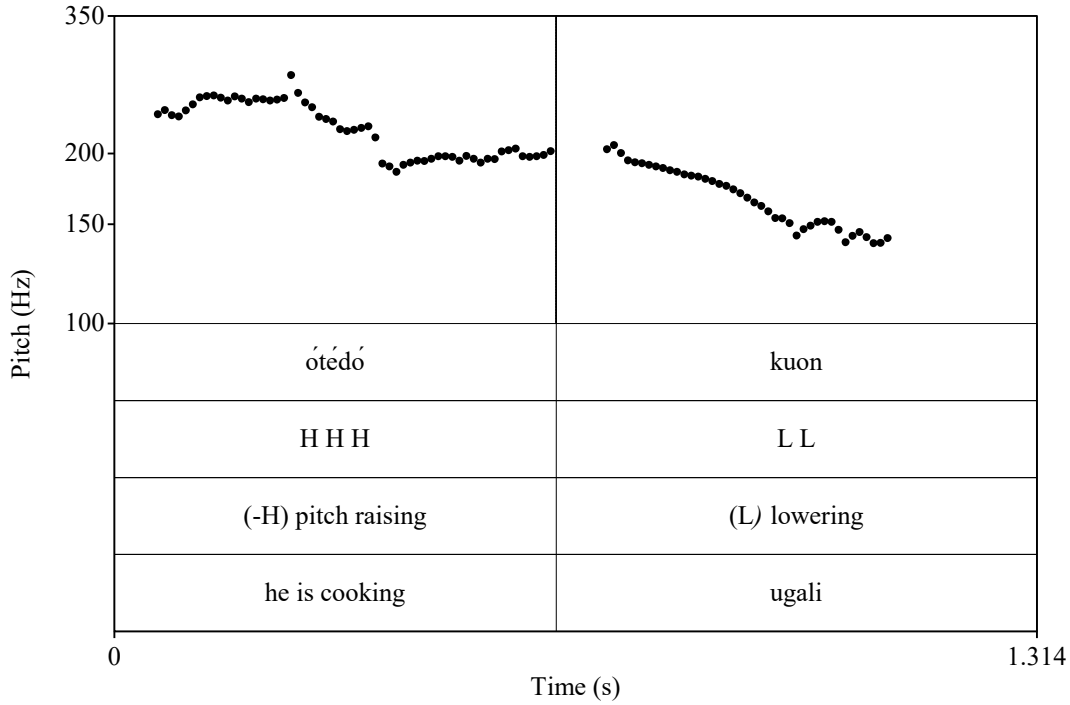


Figure 6.4 Pitch contour for the answer in (79) with pitch raising on the pre-focused word, produced by speaker SY

In this figure, it is noted that the focused word *kuon* 'ugali' lacks pitch prominence, although this might be due to the word being sentence-final and thus it is affected by boundary L%. The pre-focused word is raised, as it is produced above 230Hz. With examples (78) and (79) it seems contrastive focus has some sort of pitch raising effect, however, the exact domain for

the raising needs further research. This is because when focused words occur clause-finally they are normally affected by the boundary L%. Also, it might be difficult to determine the pitch effect if the focused word has lexical H-tones as in (78). Therefore, more data that will include longer sentences with focused words in different positions as well as different tonal patterns are needed to confirm whether the pitch raising noted above only affects the pre-focused words or can extend to the entire clause.

To conclude on focus, it is generally observed that the focused constituent is non-prominent as seen in Figures 6.1 to 6.4. In Figure 6.1, the focused constituent *wuoro* 'father' is LLL with no prominence. The word follows the pattern of the base clause only with overall pitch raising. In Figure 6.2, *káwúonó* 'today' is HHLH but ends low since it is at the end of a sentence. In Figure 6.3 *kuon* 'ugali' is LL and *túgó* 'playing' is HH but both are low with no prominence; and further when *kuon* is final in Figure 6.4 it is still low.

6.3 Clefts

The syntactic and pragmatic aspects of clefts in Luo are discussed in detail by Okoth-Okombo (1997) and Cable (2012). Generally there are various types of cleft constructions. Cable (2012) presents two types of cleft constructions possible in Luo. These are '*en XP má* ' and '*XP e má* ' which mean 'it is XP who' and 'XP is who', respectively. Cable refers to the *en* particle in the first construction as the 'normal copula' and the *e* particle in the second construction as the focus copula with *má* as a complementizer. While Cable shows that the *en*/*e* particles co-occur with a complementizer *má*, Okoth-Okombo provides examples where the focus copula *e* can occur without the complementizer. Cable points out further that the

focus copula *e* should follow a referential DP as in (80) while the normal copula *en* should precede the DP as in (81). In this case, the reverse is ungrammatical as shown in (82) and (83).

80. ɲaθî e má ríémbó ndígá
child is COMP ride bicycle
'It is the child who rides a bicycle.'

81. en ɲaθî má ríémbó ndígá
is child COMP ride bicycle
'It is the child who rides a bicycle.'

82. * ɲaθî en má ríémbó ndígá
child is COMP ride bicycle
'The child is who rides a bicycle.'

83. * e ɲaθî má ríémbó ndígá
is child COMP ride bicycle
'The child is who rides a bicycle.'

For the case of non-referential DPs, such as wh-words, Cable (2012) argues further that the DP must be preceded by a normal copula *en* as in (84). Meanwhile, the focus copula *e* is never used with non-referential DPs. Thus, the structure '**ŋa e*' is ungrammatical (see example 85). Therefore, the general rule is that, for referential DPs, the normal copula *en* precedes the DP while the focus copula *e* follows the DP, and the reverse is ungrammatical.

For non referential DPs only the normal copula *en* can precede the DP. This is because wh-questions in Luo are in situ and they can only be fronted using the '*en XP má'.*

84. en ŋa má ríémbó ndígá
is who COMP ride bicycle
'Who is it that rides a bicycle?'

85. *ŋa e má ríémbó ndígá
who is COMP ride bicycle
'Who is it that rides a bicycle?'

Another thing to note about clefts is that apart from the subject DPs already shown above, other categories can also occupy the XP position. This is shown in (86) and (87) for objects and locatives, respectively.

86. pi e má ɲaθî móðó
water is COMP child want
'Water is what the child drinks'

87. e kiɟîɟi e má i-riêmbe ndígá
LOC village is COMP LOC-ride bicycle
'In the village is where a bicycle is ridden.'

Doetjes *et al.* (2004) in their analysis of clefts in French refer to the wh-particle which corresponds to a relative clause (herein referred to as a complementizer clause) as a CODA. I

will adopt this term and meaning in this chapter. Turning to the prosodic realization of clefts, it is observed that the XP has an expanded pitch while the CODA has a compressed pitch. This is illustrated in Figures 6.5 and 6.6 below. Figure 6.5 is an illustration of sentence (80), where the DP is composed of a L HL L tonal pattern. The Ls on the DP, as observed, are produced at a higher pitch register, while the CODA which is mainly composed of H-tones is compressed. Similarly, Figure 6.6, which is an illustration of sentence (87), shows high pitch register on the fronted PP. In this case, the L LHL L tonal pattern is produced above 200Hz while the CODA is compressed. The pitch expansion on the XP indicates that there is a -H boundary at the left edge of the XP which affects the entire phrase, making it intonationally distinct from the CODA.

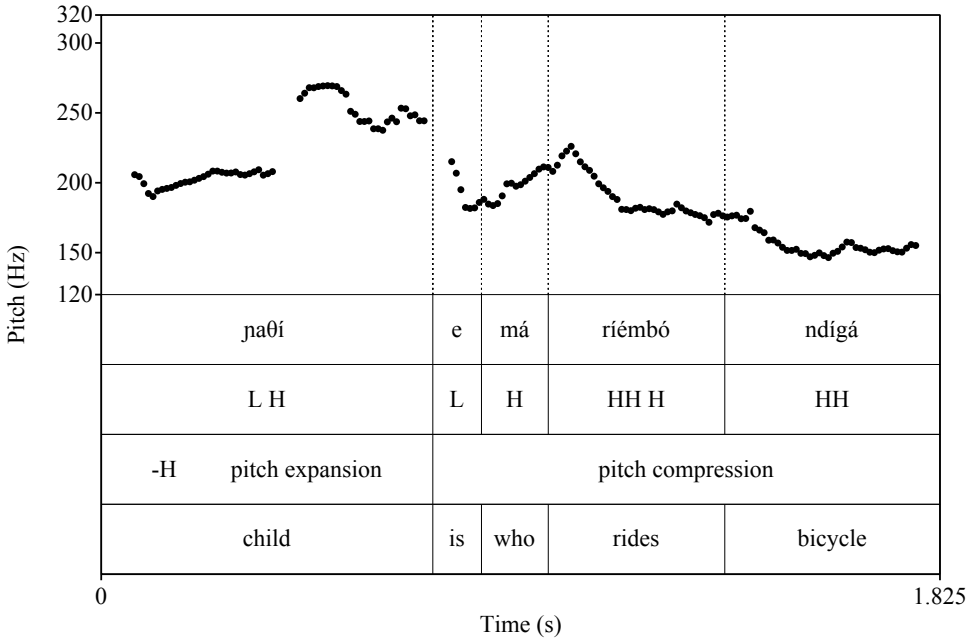


Figure 6.5 F0 contour for sentence (80) with pitch expansion on the subject DP and compression on the CODA, produced by speaker SY

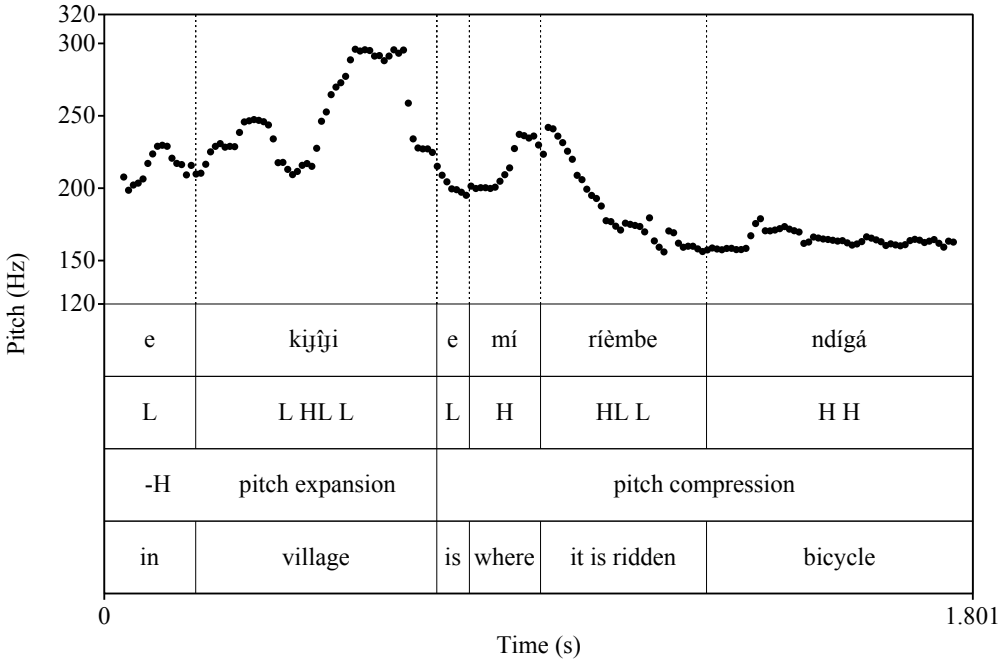


Figure 6.6 F0 contour for sentence (87) with pitch expansion on the fronted PP and compression on the CODA, produced by speaker SY

6.4 Dislocations

Dislocations are constituents which are syntactically separated from the main sentence and if omitted the sentence remains grammatical. Subjects and objects are more likely to be dislocated either sentence-initially or sentence-finally being associated with left and right dislocations, respectively. Other categories such as adverbials can also be dislocated. According to Güldemann et al. (2015), left dislocations are more common cross-linguistically than right dislocations. In Luo, subjects, objects and adverbs can be right- or left-dislocated as analysed in the following sub-sections.

6.4.1 Right dislocations

Right dislocation constructions involve a clause which is linearly followed by an XP at the right periphery. The dislocated constituent can be a subject, object or adverb. According to Truckenbrodt (2015) as well as Ott and de Vries (2016) a dislocated constituent can be in form of backgrounding (normally called right dislocation, RD) or an afterthought (AT). The former occurs when the right-dislocated constituent has an antecedent in the preceding clause. In this case, Truckenbrodt (2015) and Ott and de Vries (2016) argue that the host-internal pronoun resumes a discourse topic at the right periphery. Meanwhile, an afterthought variety occurs when the right-dislocated constituent introduces discourse-new information. In this case, there is no antecedent in the preceding clause. In this chapter I will focus on the backgrounding variety, as it is the common occurence.

Regarding subject dislocation in Luo, Cable (2012) argues that it is possible to have subjects in the post-verbal position in passives but not in active sentences. He gives the examples in (88) and (89) showing that (88) is acceptable while (89) is not. These examples from Cable (2012: 657) show that syntactically there are restrictions on moving subjects rightwards.

88. Ne ok one Onyango gi Ochieng
PST NEG PASS.see Onyango by Ochieng
'Onyango wasn't seen by Ochieng.'

89. *Ne ok o-neno Ochieng, Onyango
PST NEG 3SG.see Ochieng, Onyango
'He didn't see Ochieng', Onyango.'

Although clumsy as argued by Cable, native speakers think having post-verbal subjects can be tolerated in certain contexts. For example, it is possible to have a sentence like (89) when the interlocutors do not share background information about the referent and then the speaker is prompted to resume the discourse topic, making the subject appear in the post-verbal position even in an active voice. This is further exemplified in (90) and (91). (90a) and (91a) show the subject in its canonical pre-verbal position. In this case the subject does not co-occur with its pronominal marker unless it is topicalized. Subject topicalization is discussed below in sub-section 6.4.3. The constructions in (90b) and (91b) show the subject occurring post-verbally as a right dislocation. As argued by Cheng and Downing (2009), information structure, such as right dislocation, can favour non-canonical word order. As already mentioned above, for the subject to be dislocated to the right it requires a referent on the verb. If such a marker is not present then the sentence becomes ungrammatical. In (90b) and (91b) the pronominal forms are the prefixes *gi-* and *o-*, respectively.

90. (a) ***Canonical order***

món	o-têdo	kuon	piyo
women	PRF.cook	ugali	quickly

'The women have cooked ugali quickly.'

(b) ***Right dislocation***

gi-têdo		kuon	piyo,	món
SM3PL-PRF.cook		ugali	quickly	women

'They have cooked ugali quickly, the women.'

91. (a) ***Canonical order***

wuoro míyó ɟi kuon

father PROG.give people ugali

'Father is giving people ugali.'

(b) ***Right dislocation***

ó-míyó ɟi kuon, wuoro

SM3SG-PROG.give people ugali father

'He is giving people ugali, father.'

Objects can also be right-dislocated so long as there is a pronominal object marker on the verb. In (92) the object is marked by the suffix *–gi* (see chapter 1 section 1.3 for an overview of Luo pronouns).

92. ɲiθîndo ó-sé-ŋíéwó-gí, réc

children PRF-already-buy-3PL fish

'The children have already bought them, the fish.'

The prosodic profile of right dislocations shows that the dislocated items are prosodically phrased together with the main clause, regardless of the fact that they are morphologically marked in the clause. This is illustrated in Figures 6.7 to 6.9. In Figure 6.7, the dislocated subject *món* 'women' is affected by downstep as it would any H-toned word clause finally. Likewise, in Figure 6.8 the H-toned dislocated object *réc* 'fish' is produced at relatively the same pitch register as the previous Hs in the main clause. In Figure 6.9, the subject *wuoro*

'father' is also affected by declination by showing a gradual fall in pitch as in any L-tone sequence.

One can argue that the right dislocation is compressed and hence has a –L boundary at the left edge. For this to be possible for Luo, it is expected that the final syllable(s) would exhibit some sort of final effect. Looking at Figure 6.7, for example, the L-toned final word *piyo* 'quickly' in the host-clause does not show a drastic final lowering effect but rather it has a more flattened F0 contour, indicating a non-final effect. Similarly, in Figure 6.9, the final lowering effect is seen on the final syllable of the right-dislocated word *wuoro* 'father' instead of *kuon* 'ugali'. For the case of H-toned words, we see in Figure 6.8 that all the H-toned words (including the right dislocation) are produced within the same pitch range. This shows that there is no intonation phrase boundary between the main clause and the dislocated XP.

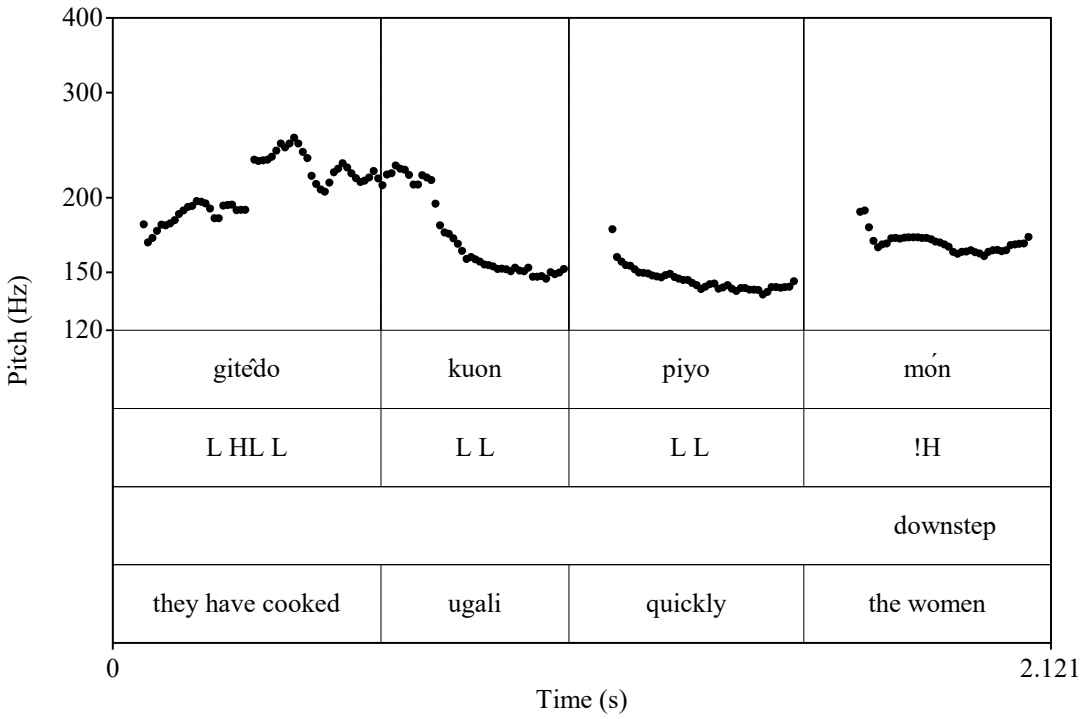


Figure 6.7 F0 contour showing a right-dislocated subject (see example 90b) affected by downstep

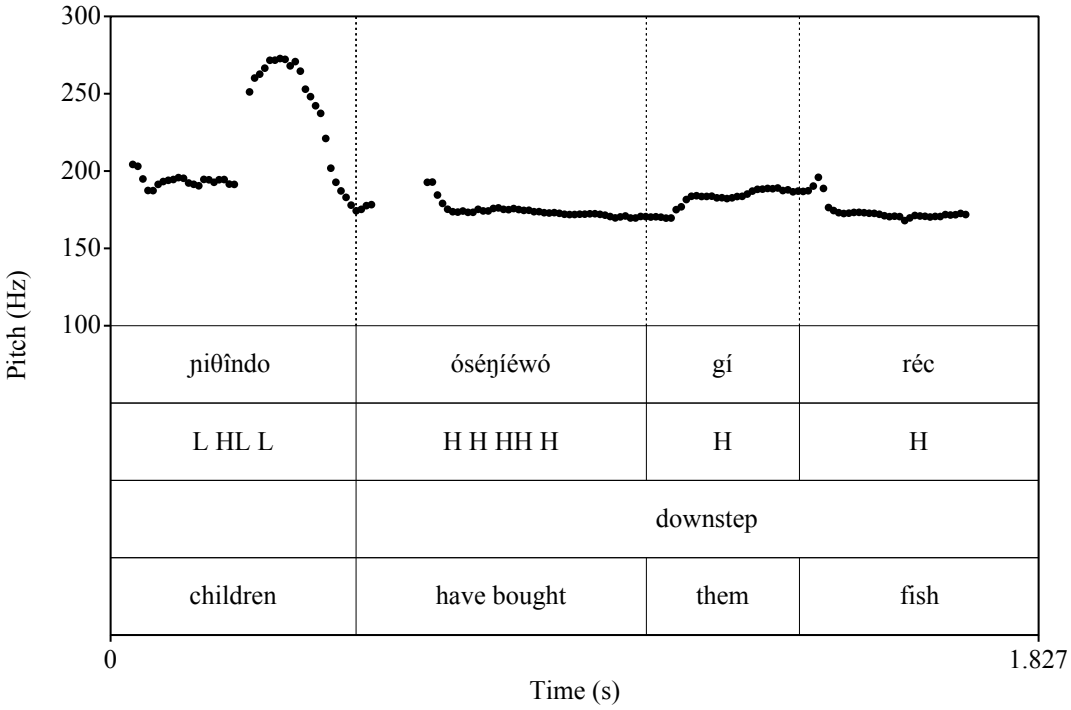


Figure 6.8 F0 contour showing a right-dislocated object (see example 92) affected by downstep

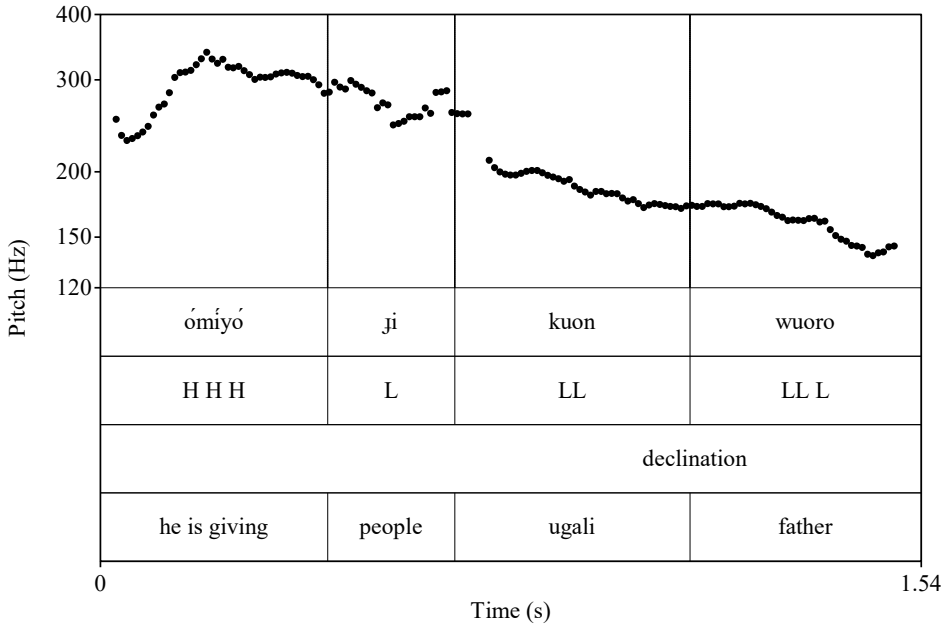


Figure 6.9 F0 contour showing a right-dislocated subject (see example 91b) affected by declination

The lack of an intonation boundary between the two juxtaposed structures shows that right dislocations can be morphosyntactically marked as they belong to two separate clauses (i.e CP_1 and CP_2) but they are not prosodically marked. This case is well-argued by Truckenbrodt (2015) and Ott and de Vries (2016). In their analyses of German and Dutch they argue that in backgrounded dislocation (where the correlate resumes a discourse topic), the XP is deaccented and realized with a low level intonation. Meanwhile, in the afterthought dislocation (where there is an introduction of discourse-new information) the XP is accented and realized with a focal stress. This means the backgrounded dislocations are intergrated into the intonation phrases of their host clauses despite being two distinct syntactic constituencies. Adding to this, Truckenbrodt (2015) argues that the backgrounded XPs and their host clauses together form a single speech act and hence lack prosodic boundaries. This

is contrary to Selkirk's (2011) Match Theory as it shows a mis-match between syntactic and prosodic constituents.

There are two main types of prosodic phrasing proposed for the dislocated constructions. In Bantu languages it is pointed out that an asymmetric phrasing pattern occurs when right dislocations are phrased separately from the main clause but left dislocations are phrased with the main clause, and symmetric phrasing occurs when both right and left dislocations are phrased separately (Cheng & Downing, 2009; Downing, 2011). Another proposed pattern regards left dislocations which are said to be phrased with the main clause or not (ibid). In Zulu, as well as in other Bantu languages, Cheng and Downing report a constistent phrasal break before the right dislocation, the case which is not observed in Luo. The kind of asymmetry observed in the Luo data presented above mirrors the proposed pattern by having right dislocations phrased together with the main clause while left dislocations are phrased separately. We return to left dislocation below.

6.4.2 Left dislocations

In this section we will consider left dislocation for objects. We return to left-dislocated subjects in section 6.4.3 below. Generally, in left dislocation, constituents are moved clause-initially as topics. Hence left dislocation has been used interchangeably with topicalization. Ross (1967) as well as Rivero (1980) differentiate the two terms by arguing that the former applies a copying rule whereby the left-dislocated NP has a pronoun which is anaphorically related to it in the main clause. Topicalization, on the other hand, has a chopping effect whereby the topicalized NP leaves a gap in the original place. Feldhausen (2016) considers left dislocation with resumptive pronouns as hanging topic left dislocation (or clitic left

dislocation, in case a weak pronoun is used) and those without resumptive pronouns simply as left dislocations. The former can also be introduced with '*As for'* structure (see example (93a) below). In this chapter, I will generally use the term left dislocation with and without resumptive pronouns to respectively refer to those with and without object markers in the main clause.

In Luo, post-verbal XPs can be left-dislocated with an anaphoric referent in the main clause, as seen in (93a). Here the dislocated object *ɲiθîndo* 'children' is marked by the suffix *-gí* 'them'. Alternatively, the left-dislocated XP can occur as a topic even without the object marker as in (93b).

93. (a) ɲiθîndo miyo bíró luoko-gí káwúònó goðíàmbo
children mother will wash-3PL.OBJ today evening
'(As for) the children, mother will wash them today in the evening.'

(b) ɲiθîndo miyo bíró luoko káwúònó goðíàmbo
children mother will wash today evening
Children, mother will wash today in the evening.'

Further cases of left dislocation without resumptive pronouns are presented in (94a) and (95a), but they can also occur with their respective object markers, in this case the suffix *–e* (see examples (94b) and (95b)).

94. (a) kwac, ŋiθîndo négó
leopard children PROG.kill
'The leopard, children are killing.'

(b) kwac, ŋiθîndo nég-é
leopard children kill-3SG.OBJ
'The leopard, children are killing it.'

95. (a) ŋuka, ɟatedo núdó gí oluth-kuon e ɟikôn
porridge cook PROG.cook with cooking-stick in kitchen
'Porridge, the cook is cooking with a stick in the kitchen.'

(b) ŋuka, ɟatedo núd-é gí oluth-kuon e ɟikôn
porridge cook cook-3SG.OBJ with cooking-stick in kitchen
'Porridge, the cook is cooking it with a stick in the kitchen.'

The prosodic contours for the left dislocations are illustrated in Figures 6.10 to 6.12. These figures show that there is a prosodic boundary at the right edge of left dislocation. This is indicated by a boundary L% showing a separation between the dislocated XP and the following main clause. Additionally, the main clause in all the cases presented are marked at the left edge by pitch reset, whereby the L-tones of the initial syllables reset back to 200Hz. The pitch reset can be preceded by a pause as clearly noticed in Figures 6.10 and 6.11. But pauses are not mandatory as noted in Figure 6.12.

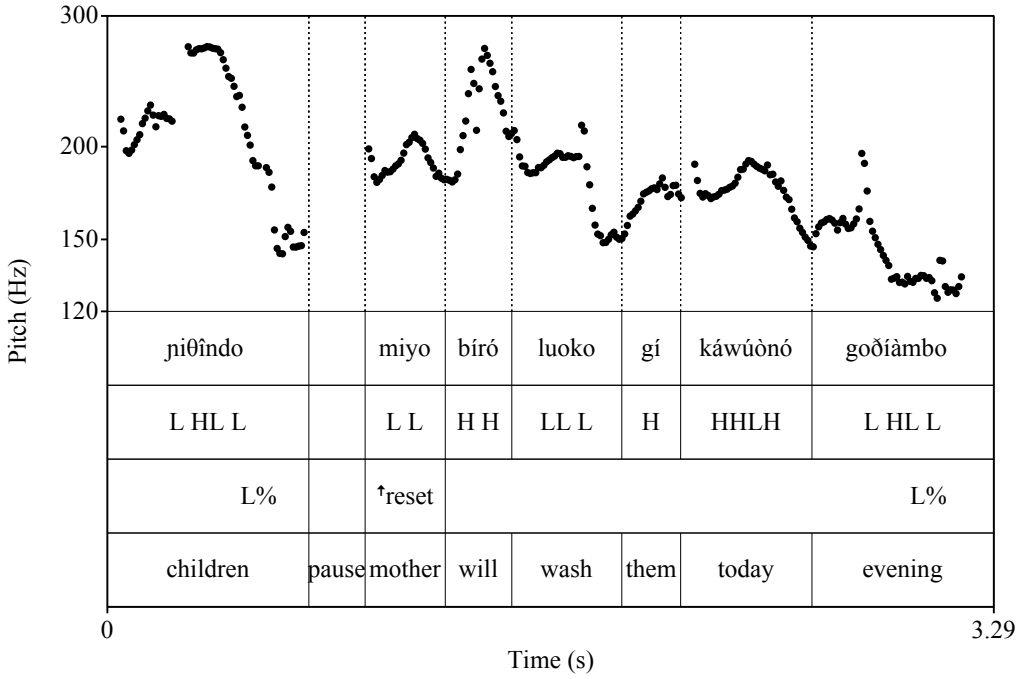


Figure 6.10 F0 contour for sentence (93a) showing a boundary L% at the right edge of the left dislocated XP constituent followed by a pause and pitch reset on the preceding clause

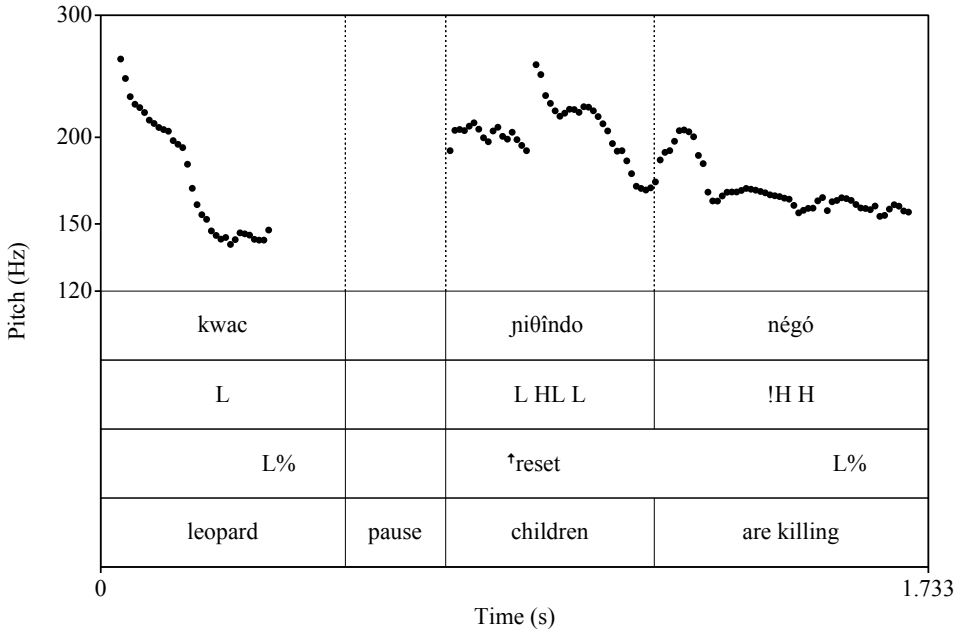


Figure 6.11 F0 contour for sentence (94a) showing a boundary L% at the right edge of a topicalized constituent followed by a pause and pitch reset

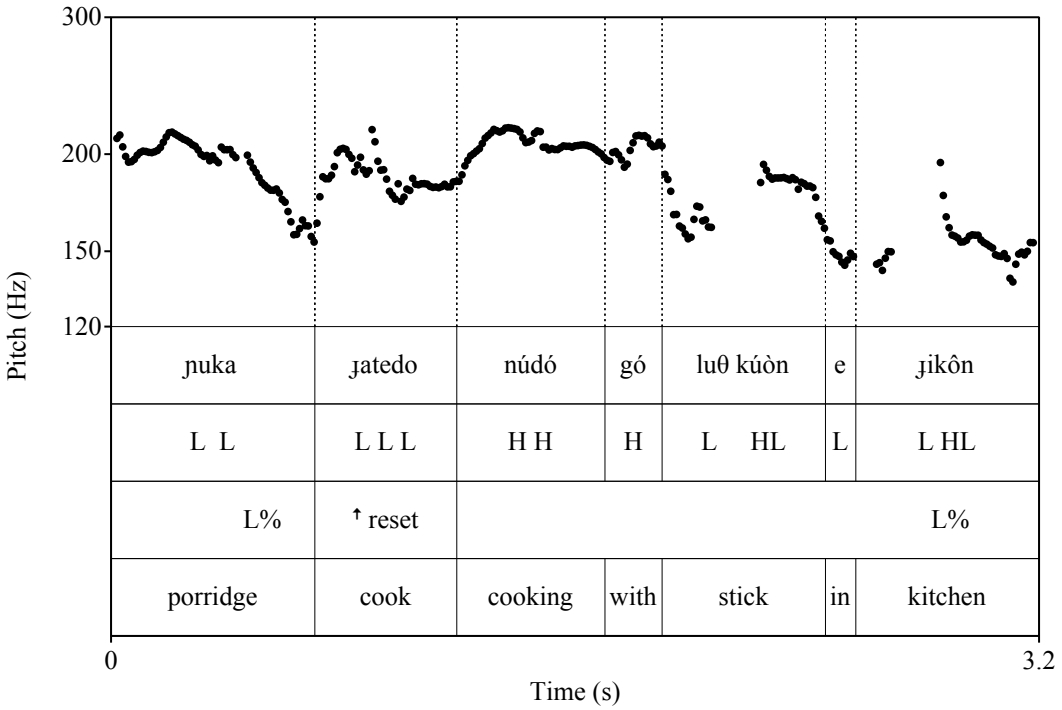


Figure 6.12 F0 contour for sentence (95a) showing a boundary L% at the right edge of the left dislocation (without resumptive pronoun) and pitch reset (without a pause) at the left edge of the main clause

As opposed to right dislocated XPs discussed above, left dislocated XPs seem to be phrased separately from the main clause. Furthermore, it is noted that there is no prosodic difference with respect to the right edge between constructions with resumptive pronouns and those without pronouns. Feldhausen (2016) also reports similar results for Spanish left dislocations. He argues that left dislocations with or without resumptive pronouns have similar intonational patterns, both realized by a rising nuclear pitch. These results also concur with what is reported in many African languages (see Downing, 2011 for an extensive study on dislocations in Bantu languages) with left dislocations as separate entities. In Luo, low boundary L% (in the form of final lowering or downstepped H) is the primary intonation cue for right edge. This is not only noted in left dislocation but also in other constructions like

declaratives and questions. A pause can be considered as a secondary cue which can be present or not.

6.4.3 Topicalized subjects

In this section we consider subjects occurring as topics at the left periphery. In principle, lexical subjects in Luo do not co-occur with their pronominal markers. This occurence is only possible when the lexical subject is a sentence-initial topic. The syntactic difference between a sentence with and without a topicalized subject is shown in (96) and (97). (96a) contains a sentence with a lexical subject *wuoro* 'father' which co-occurs with the subject marker *o-* while (96b) has no such marker. This makes the lexical subject in (96a) a topic. The same is noted in (97a) and (97b) where the lexical subject *ɲiθîndo* 'children' is marked by *gi-* on the verb.

96. (a) wuoro, ó-bíró ði luw-o réc goðíàmbo
father SM 3SG-will go catch-INF fish evening
'Father, he will go to catch the fish in the evening.'

(b) wuoro bíró ði luw-o réc goðíàmbo
father will go catch-INF fish evening
'Father will go to catch fish in the evening.'

97. (a) ɲiθîndo, gí-bíró tugo e kúóyó kîɲ godiocíèŋ
children SM3PL-will play in sand tomorrow afternoon
'Children, they will be playing in the sand tomorrow afternoon.'

(b)	ɲiθîndo	bíró	tugo	e	kúóyó	kîɲ	godiocíèŋ
	children	will	play	in	sand	tomorrow	afternoon

'The children will be playing in the sand tomorrow afternoon.'

Topicalized subjects syntactically behave like left-dislocated objects, as they are also clause-external. According to Cheng and Downing (2009) since subjects occur in a similar syntactic position as other topics in the left periphery, they should also have similar prosodic features as other topicalized constituents. Thus, as it was to left-dislocated objects above, topicalized subjects also have boundary L% at the right edge and pitch reset at the left edge of the main clause. Additionally, a pause between a topicalized subject and the rest of the sentence can also signal a prosodic break. All these cues are illustrated in Figure 6.13.

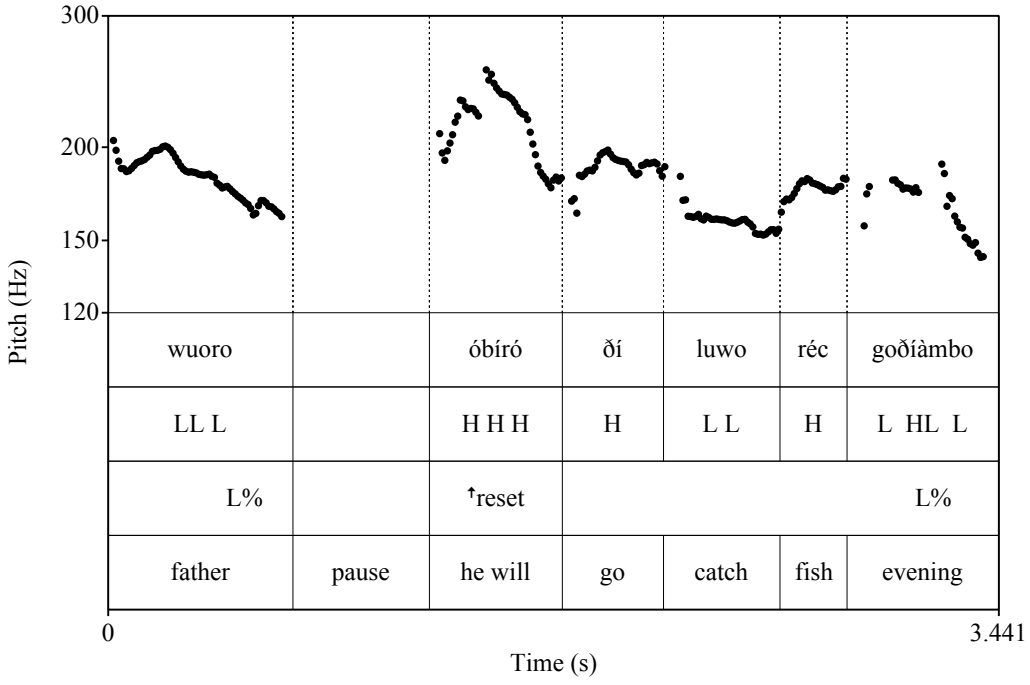


Figure 6.13 F0 contour for sentence (96a) showing a boundary L% at the right edge of a topicalized subject and pitch reset at the left edge of the main clause

For the non-topicalized subjects, there is no boundary L% at the right edge of the subject, as noted in Figure 6.13. This indicates that non-topicalized subjects in Luo are phrased together with the following constituent, in this case with a VP.

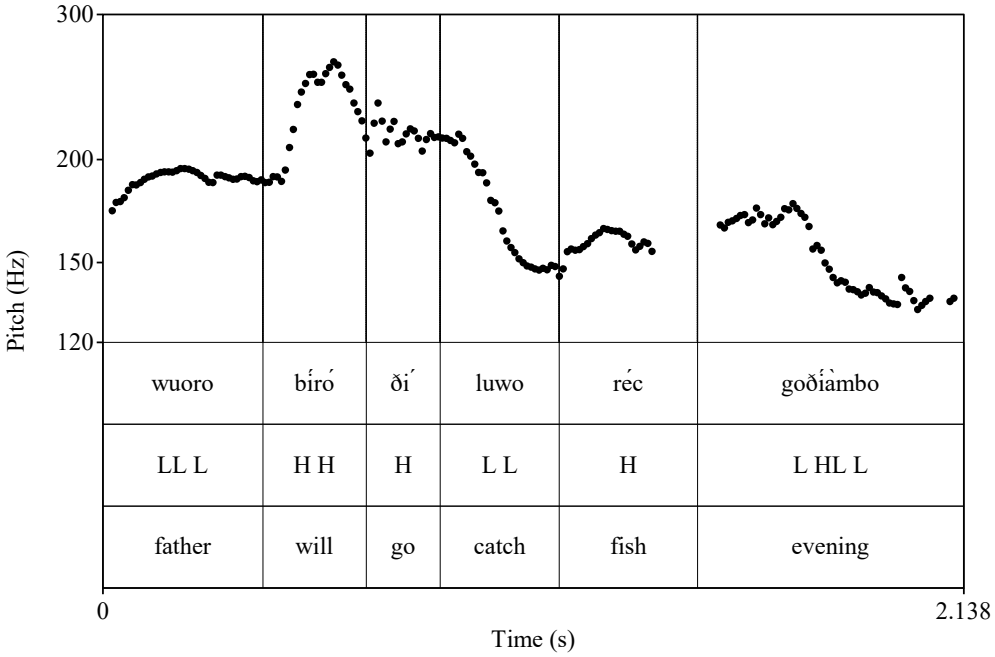


Figure 6.14 F0 contour for sentence (96b) without prosodic break at the right edge of a non-topicalized subject

To sum up on the prosody of the left periphery, it is clearly noted that topicalized objects and subjects have similar intonational patterns, realized by boundary L% at the right edge and may be followed by a pause. For the objects, the features occur whether the object has a resumptive pronoun or not. This is different for topicalized subjects which, without resumtive pronouns are not considered to be topics.

6.5 Conclusion

In this chapter an analysis of the syntactic and prosodic aspects of information structuring in Luo is given. The first aspect covered is focus and information structure of wh-questions (discourse-new information) or presentational focus, contrastive focus and clefts. As in other tone languages, Luo lacks prosodic prominence on the focused constituent but rather shows an overall pitch range expansion for non-contrastive or presentational focus. This is phonologically marked by a -H boundary at the left edge of the clause which extends to the entire clause. Contrastive focus is also marked by -H boundary at the left edge but its extension seems to be restricted to the pre-focused constituent. However, this aspect needs further investigation. It should be noted that -H boundary at the left edge is also observed in questions. Clefts, on the other hand, have pitch raising on the fronted XP while the following part (i.e. the CODA) is compressed.

Lack of focus prominence is reported in many African languages contrary to what has been reported in intonation languages which generally have pitch accentuation on the focused constituent (Gussenhoven, 2004). In African languages the presence of focus prosody is still questioned. In the recent collection by Downing and Rialland (2016) almost all of the African languages studied lack prominence on the focused constituents. In some languages, particularly the Bantu languages, prosodic phrasing is the main indicator of focus.

Apart from focus, this chapter has also investigated the syntactic and prosodic aspects of dislocations. Here three types of dislocation have been analysed: right dislocation, left dislocation and topicalized subjects. It is observed that left and right dislocations are

morphosyntacticaly marked, as they form a separate constituent structure from the main clause. Prosodically, we have seen that right dislocations are phrased together with the host clause while left dislocations, including topicalized subjects, form separate prosodic units. This is because, for the right dislocation there is no evidence of a prosodic break or phrase final intonation effects between the main clause and the dislocated constituent. But, for the left dislocation, there are intonational effects observed. These are boundary L% at the right edge of the dislocation and pitch reset at the left edges of the main clause. This kind of phrasing shows asymmetry in Luo which contrasts with what is reported in other African languages, particularly the Bantu languages.

Chapter Seven

Intonation of Complex Sentences

7.0 Introduction

This chapter presents the intonation structure of complex sentences in Luo where two types of complex structures are investigated. This includes those with subordinate conjunctions and those with coordinate conjunctions. It is generally agreed that syntactic clauses roughly align with intonation phrases (IPs), which form the higher level of prosodic phrasing following the prosodic hierarchy by Selkirk (1984). Since complex sentences are made up of more than one clause, the assumption is that they prosodically also involve recursive intonation structures (Selkirk, 2009). The phonetic correlates that cue such intonation phrasing differ cross-linguistically, but it is commonly noted that pitch reset (full or partial) or a register high tone is a main cue at the left edge while boundary tones, as well as pauses, function as the main cues at the right edge. Additionally, it is observed that global effects such as pitch register expansion (PRE) or pitch register compression (PRC) can also signal complex sentences.

In this chapter, intonational effects of various complex structures in Luo are examined to evaluate whether the assumed recursive structures also apply. It is noted that both local and global effects as outlined above are also observed in Luo complex structures. These effects are systematically observed in sentences with subordinate clauses as well as those with coordinate clauses. We can however observe that the complementizer *ni* does not have any intonation marking at the left or right edge of an embedded clause or main clause, making it questionable whether all complex structures are intonationally recursive.

The chapter is organized as follows: Section 7.1 concentrates on complex sentences with subordinate clauses. Three types of subordinate clauses are investigated: relative clauses, adverbial clauses and complementizer clauses. Section 7.2 concentrates on complex sentences with independent clauses, also known as coordinated clauses. Section 7.3 concludes the chapter.

7.1 Subordination

Subordination involves syntactic dependency between two clauses. Thus, a subordinate clause is embedded to a superordinate or matrix clause (Fabricius-Hansen & Ramm, 2004). According to Cristofaro (2003) subordinate clauses under traditional classification include those introduced by a complementizer like 'that', relative clauses and adverbial clauses. In Luo, such clauses are introduced by the complementizer *ni* 'that', relativizer *má* 'who/which and adverbial conjuncts such as *ka* 'when/if', *nikec* 'because', etc. These are further analysed in the following sections.

7.1.1 Relative clauses

The first type of subordinate clause to be investigated is a relative clause. In Luo, both restrictive and non-restrictive relative clauses are marked by the relativiser *má.* Normally, the relativiser *má* appears in a sentence as an independent word. However, in some cases, the vowel /a/ is deleted when it is followed by another vowel, and hence the remaining *m*-, as well as its floating H tone, reattaches to the following word, making the relative marker appear as a prefixed morpheme. The examples in (98) and (99) show the relativiser *má* occurring as an underlyingly separate word in the sequences *má ocûŋ* and *má ondîko*. In

connected speech they surface as *mócûŋ*, and *móndîko*, respectively.[15] The restrictive relative clause in (95) identifies which boy the speaker is referring to, in the context when there is more than one boy known by the participants. By contrast, the non-restrictive clause in (99) adds supplementary information about teacher Juma.

98. wúóyí má o-cûŋ e báθ mêsa en ɲaθí-ná
boy REL 3SG-stand beside table is child-1SG.POSS
'The boy who is standing beside the table is my son.'

99. ɟapúóɲɟ ɟúma, má o-ndîko buk, kéc ahíɲá
teacher ɟúma, REL 3SG.write book strict very
'Teacher Juma, who has written the book, is very strict.'

Let us begin with the restrictive relative case in (98). Syntactically, it is assumed that this sentence is made up of two clauses: the matrix clause which contains the relative clause *wúóyí má o-cûŋ e báθ mêsa* and the main clause *en ɲaθíná*. These clauses comprise smaller constituents such as the NP *wúóyí*; Complementizer *mócûŋ*; PP *e báθ mêsa*, and VP *en ɲaθíná*. The question which arises here is whether syntactic phrasing is isomorphic to

[15] There are cases where the vowel in *má* is not deleted even when it is followed by another vowel, as in:
ɟapúóɲɟ *má* od-é ber
Teacher REL house-3SG.POSS good
'The teacher whose house is good…'
The reasons for not deleting the vowel in this case is not very clear but the general observation is that when nouns follow *má* deletion is blocked, meaning nouns do not allow *m-* to be prefixed on them.

prosodic phrasing. To answer this question let us examine the prosodic correlates from the pitch contour presented in Figure 7.1 below.

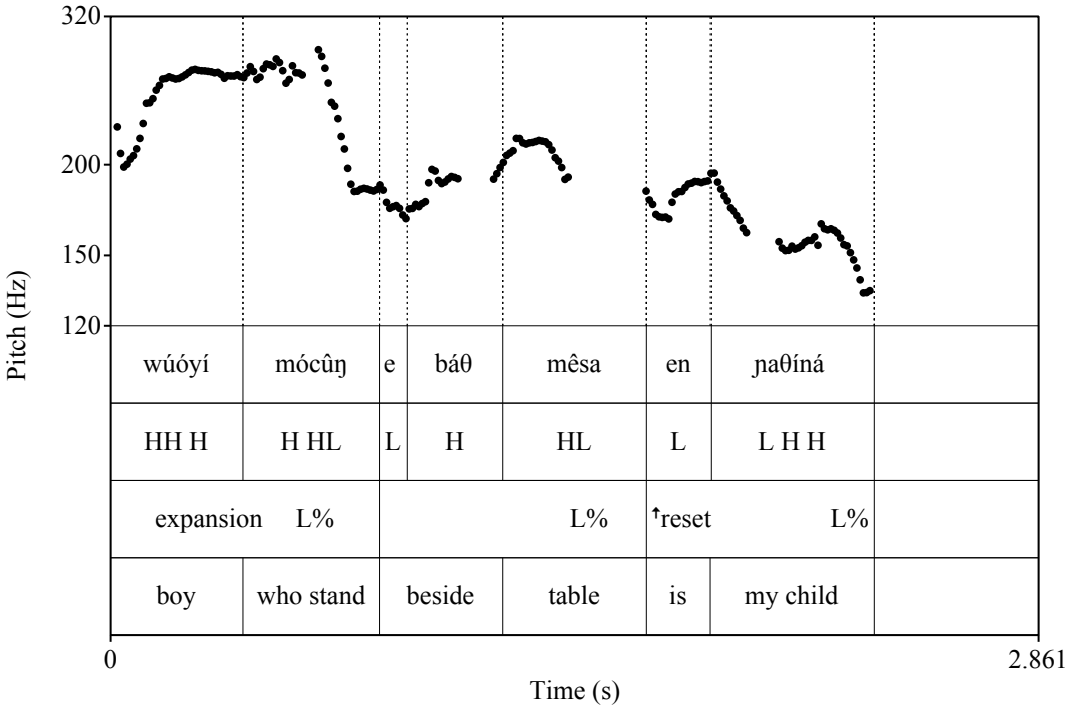


Figure 7.1 F0 contour of a complex sentence with a restrictive relative clause (98) showing PRE on the complex NP with a boundary L% at the right edges of the NP, PP and VP and pitch reset at the left edge of the VP

In this figure, it is observed that the head noun (NP) *wúóyí* and the complementizer *mócûŋ*, which together form a complex NP *wúóyí mócûŋ* 'a boy who is standing' have expanded pitch register (PRE). From the previous discussion it is noted that generally L-tones are produced below 200Hz while H-tones are produced at or slightly above 200Hz. In this figure, however, the H-tones in the complex NP have a more expanded register, between 250Hz and 300Hz. This shows that there is a boundary –H tone at the left edge of the complex NP, which, however, does not extend to other constituents. There is also a prosodic effect at the

right edge of the NP noted on the final syllable *-cûŋ*. This syllable has a dramatic fall in pitch, going below 200Hz as an indication of boundary L% at this edge. This suggests that the complex NP *wúóyí mócûŋ* 'a boy who is standing' forms a separate prosodic phrase from the following constituent, in this case the PP.

The PP *e báθ mêsa* 'beside the table' and the VP *en ɲaθíná* 'is my child' are also prosodically marked at their right and left edges. Notice that the right edge of the PP is also cued by a boundary L% in Figure 7.1. Among the four speakers, three exhibited L% at this edge and only one speaker (i.e. SY) exhibited a slight rise in pitch. Having boundary tones at this edge also suggests the presence of an independent prosodic phrase for the PP. The VP, which in this case is the main clause of the sentence, is cued at the left edge by a pitch reset, which is realized by a LH (rising) contour. This is evidenced by a prosodic break between the two L-tones, that is, the L-tone at the final syllable of the PP and the one at the initial syllable of the VP. While the former shows a lowering effect phrase finally, the latter resets to 200Hz, the pitch range which is relatively similar to that of the initial syllable of the first phrase, the complex NP. Thus, the prosodic phrasing of a restrictive relative clause described above can be schematized as in (100), where IP stands for a major intonational phrase, ip for a minor intonational phrase, % for boundary tones and an upward arrow (↑) for pitch reset.

100. [$_{IP}$ [$_{IP}$ [$_{ip}$ wúóyí mócûŋ $_{L\%}$] [$_{ip}$ e băθ mêsa $_{L\%}$]] [$_{IP}$ ↑en ɲaθí-ná $_{L\%}$]]]

[$_{IP}$ [$_{IP}$ [$_{ip}$ The boy who is standing $_{L\%}$] [$_{ip}$ beside the table $_{L\%}$]] [$_{IP}$ ↑is my son $_{L\%}$]]]

With respect to prosodic phrasing, restrictive relative clauses are recursive, being associated with minor and major intonational phrases. As shown in (100), there are three distinct intonational phrases, which are realized at their right edges with boundary L% and/or with pitch reset at their left edges. First, the head noun and the restrictive relative clause are phrased together as a minor intonational phrase (henceforth, *ip*) with a boundary L%. The following prepositional phrase (PP), which also forms an *ip* has a boundary L% but no pitch reset at the left edge. On the other hand, the VP, in this case, is associated with the major intonational phrase (henceforth, IP) being identified at both left and right edges with pitch reset and boundary L%, respectively. As already mentioned above, the head noun and the relative clause have expanded pitch register while the PP and the VP exhibit compressed pitch register (see Figure 7.1). Such intonation pattern is also observed in the sentences with adverbial clauses, as we shall see in section 7.1.2. as well as in cleft constructions as discussed in chapter 6 section 6.3. Recall that cleft constructions are also introduced by the relativiser *má* plus *e* or *en*. With regard to cleft constructions, pitch expansion is observed on the initial XP while the following constituent (i.e. CODA) is compressed.

Another type or relative clause examined is a non-restrictive relative clause. As already introduced above, syntactically, this clause is also introduced by the relativiser *má* and is said to be set off from the main clause prosodically, indicated with commas in example (99). As a rule, omitting a non-restrictive relative clause leaves the remaining sentence independent and grammatical, as the clause adds non-essential information. Thus, *ɟapúóŋɟ ɟúma kéc ahíŋá* ‘Teacher Juma is very strict’ is an independent clause and the embedded clause *móndîko buk* ‘who has written a book’ is non-essential. As we did with restrictive relatives, the non-

restrictive in (99) can also be sub-divided into XPs as follows: a head noun NP *ɟapúóŋɟ ɟúma*; CP *móndîko buk*; and VP *kéc ahíɲá.* The prosodic realization of this sentence is presented in Figure 7.2.

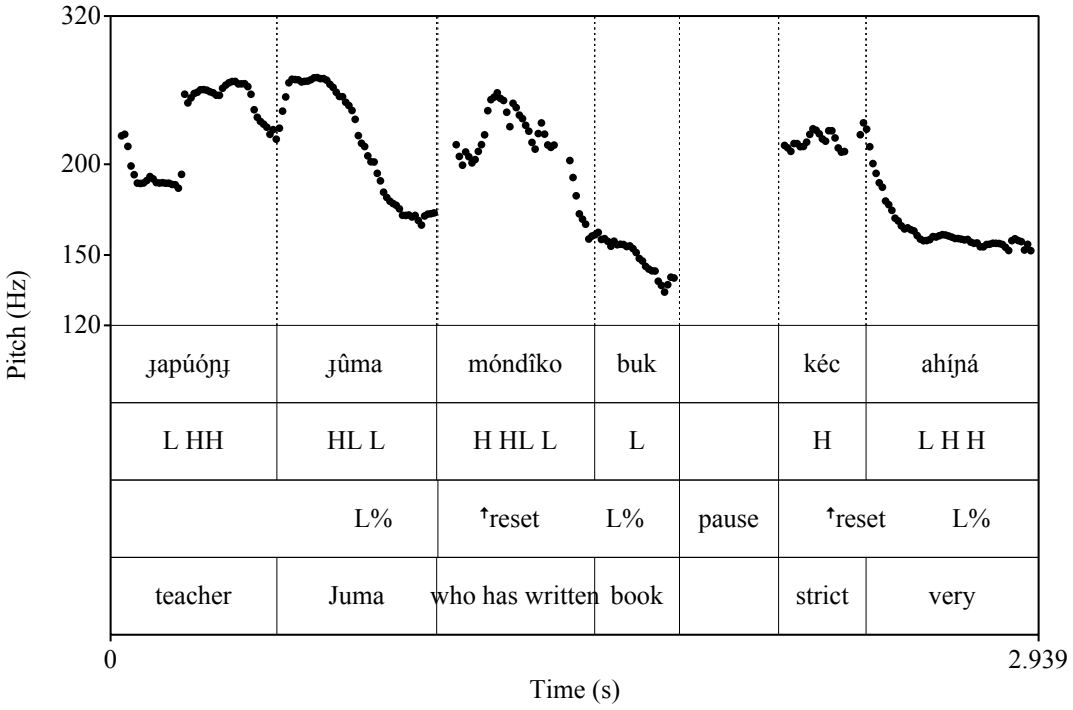


Figure 7.2 F0 contour of a complex sentence with a non-restrictive relative clause (99) showing a boundary L% preceding and following the embedded relative clause

The prosodic realization of a non-restrictive relative clause seems to be somewhat different from that of a restrictive relative clause. Firstly, there is no pitch register difference between the embedded and the main clause. All the lexical Hs (except at the boundaries) maintain the pitch range between 200Hz and 250Hz, compared to the register expansion observed in the complex NPs of the restrictive relatives, which went upto 300Hz. Secondly, the head noun is phrased separately from the non-restrictive relative clause. As observed in Figure 7.2, we see three distinct prosodic phrases realized at their left- and right edges, being associated with

major intonational phrases (IPs). In this case, the non-restrictive relative clause *móndîko buk* and the main clause *kéc ahíɲá* have pitch reset at their left edges relative to the initial pitch on the head noun *ɟapúóɲɟ ɟúma*. Additionally, each IP has a boundary L% at the right edge, whereby the final Ls are affected by final lowering and final Hs are downstepped. This intonational pattern has been consistently maintained by the three speakers. Speaker SY, on the other hand, exhibits a more sustained L-tone at the right edge of the non-restrictive relative clause rather than a real final lowering which falls to the bottom of the speaker's pitch range. Additionally, there is an audible pause realized by all speakers at the left edge of the main clause (VP). These prosodic boundaries indicate that non-restrictive relatives are also recursive. But contrary to relative clauses, non-restrictive clauses are associated with major intonational phrases (IPs), as schematized in (101) below.

101. $[_{IP}\ [_{IP}$ ɟapúóɲɟ ɟúma $_{L\%}]$ $[_{IP}$ ↑má ondîko buk $_{L\%}]$ $[_{IP}$ ↑kéc ahíɲá $_{L\%}]]$

$[_{IP}\ [_{IP}$ Teacher Juma. $_{L\%}]$ $[_{IP}$ ↑who has written the book $_{L\%}]$ $[_{IP}$ ↑is very strict $_{L\%}]]$

To sum up, both restrictive and non-restrictive relative clauses are recursive, as their edges are locally cued but with different prosodic structures. The relative clause has pitch range expansion on the head noun and compressed pitch on the following constistuents, an effect found in cleft constructions and adverbial clauses. Thus, the constituents in relative clauses are associated with ips and IPs. On the other hand, non-restrictive relatives have no pitch expansion but rather all the phrases are equally realized by a rise-fall contour, being associated with IPs.

7.1.2 Adverbial clauses

As in other languages, Luo has several adverbs which can function as adverbial conjunctions. In this section, two adverbial conjunctions are discussed. The first one is the temporal adverb *ka* which means 'when' (it can also be used as a conditional adverb 'if') and the clausal adverb *nikec* 'because'. In order to investigate the prosodic edge effects, adverbial clauses were placed both sentence-finally in their most canonical position as in (102a) and (103a), and sentence-initially as inverted orders in (102b) and (103b). Furthermore, as pointed out by Gussenhoven (2004) negation can have an impact on the prosodic structure of a sentence by reducing the excursion size and causing pitch compression in the negative construction. Thus, in order to assess whether negation in Luo has such a prosodic effect, the sentence in (103) will serve as an example. The pitch contours for these sentences are illustrated in Figures 7.3 to 7.6.

102. (a) né á-púró puoð-á ká né o-gîk
PST SM1SG-PROG.cultivate farm-1SG.POSS when PST SM3SG-PRF.arrive
'I was cultivating my farm when he arrived.'

(b) ká né o-gîk né á-púró puoð-á
when PST SM3SG-PRF.arrive PST SM1SG-cultivate farm-1SG.POSS
'When he arrived, I was cultivating my farm.'

103. (a) ok á-bí wúóyó kod-í nikec i-lêwo
NEG SM1SG-will talk with-OM2SG because SM2SG-late
'I will not talk to you because you are late.'

(b) nikec i-lêwo, ok á-bí wúóyó kod-í
because SM2SG-late NEG SM1SG-will talk with-OM2SG
'Because you are late, I will not talk to you.'

Regarding the prosodic realization of adverbial clauses, let us first consider the affirmative case in the canonical order (102a) as illustrated in Figure 7.3, where the main or independent clause is followed by a subordinate clause. First, there is expansion in pitch observed on the first clause whereby the first three H-toned syllables are produced between 200Hz and 300Hz. The following L-tones are produced below 200Hz and the final H is downstepped, as it would any H-tone at the final position. In this figure, we do not see much expansion effect on the word *puoðá.* This might be due to the word being at the final position and thus affected by final effects. Recall from the previous discussion that final lowering can affect the final syllable or a number of syllables. There is a pause which also cues a prosodic break between the main and the subordinate clause. Afterwards, there is a reduction in excursion size suggestive of pitch range compression on the subordinate clause. To elaborate on this, we see that this clause, *ká nogîk* 'when he arrived', has HLHL tonal contour produced below 200Hz when the clause appears sentence-finally. Meanwhile, the same clause when it appears sentence-initially, as we will see below in Figure 7.4, the pitch register goes above 200Hz. As global effects, pitch expansion affects the entire first clause while pitch

compression affects the entire second clause. This suggests that there is a -H boundary at the left edge of the first clause and a -L boundary at the left edge of the second clause.

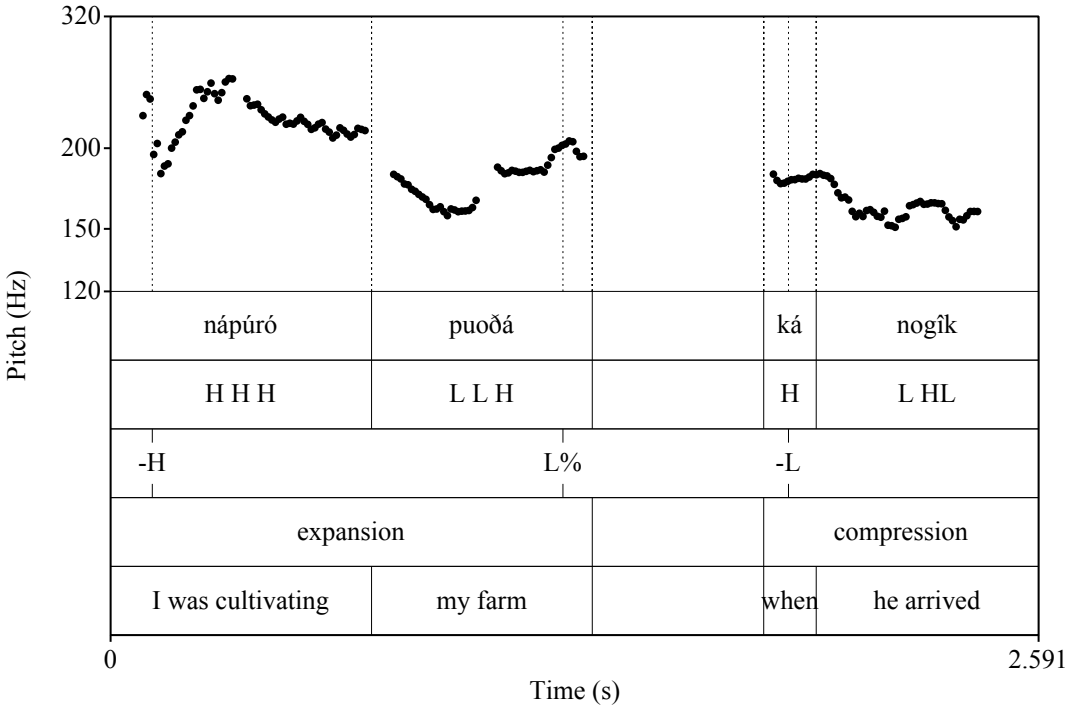


Figure 7.3 F0 contour of a complex sentence (102a) with an adverbial clause in its canonical order, showing PRE on the main clause and PRC on the adverbial clause

Figure 7.4 shows the F0 contour of the same sentence (i.e 102a) in an inverted order (see 102b). Here, the subordinate clause precedes the main clause. Like in the canonical order above, the first clause, in this case the subordinate clause *ká nogîk* 'when he arrived', has an expanded pitch register where the HLHL tonal contour is generally produced above 200Hz with the initial H targeting 240Hz and the final L lowered to 200Hz. As it was noted in the canonical order above, the inverted order is also prosodically separated by a pause, which seems to be slightly longer with the difference of 0.13sec compared to the pause in the canonical order. The main clause *nápúró puoðá* 'I was cultivating my farm' is also

compressed and produced below 200Hz regardless of the fact that the sentence begins with a sequence of H-tones. When the same clause was sentence-initial in Figure 7.3 it was produced between 200Hz and 300Hz. This, as already marked in Figure 7.4, shows a -H boundary at the left edge of the subordinate clause and a -L boundary for the main clause. In both cases (i.e. canonical and inverted) there is no pitch reset observed at the left edge of the second clause. Lack of pitch reset in this position confirms that pitch in the second clause is compressed.

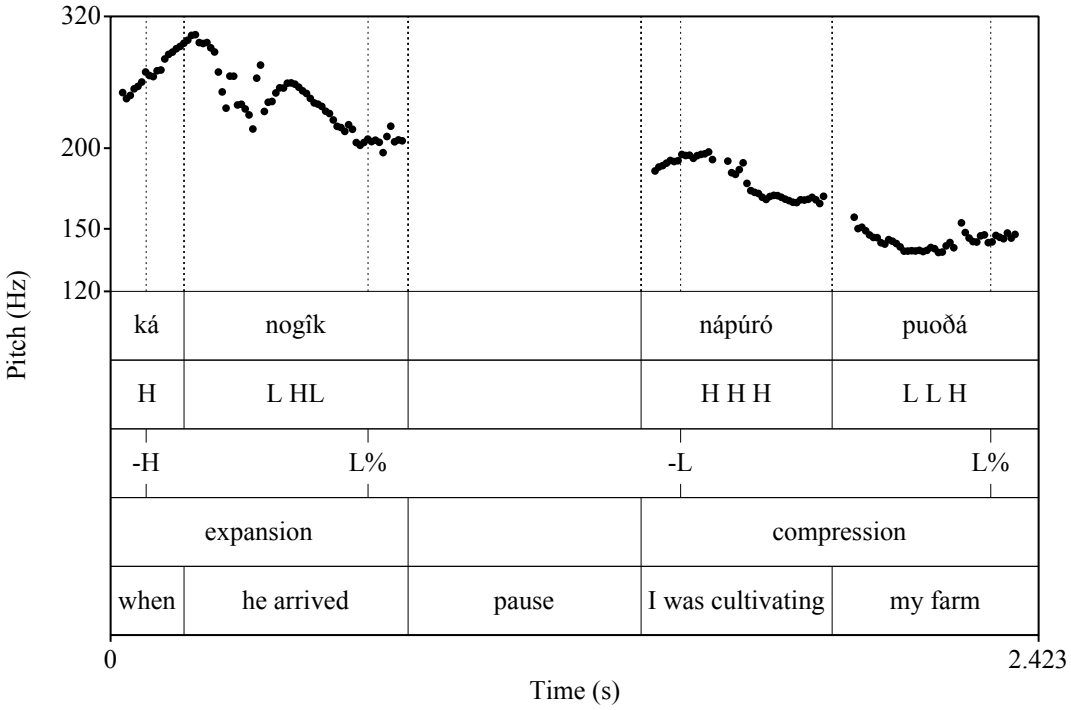


Figure 7.4 F0 contour of a complex sentence (102b) with a fronted adverbial clause, showing PRE on the adverbial clause and PRC on the main clause

Regarding the negative construction in (103), the negation word *ok* is placed on the main clause. The pitch contours for this sentence are illustrated in Figures 7.5 and 7.6 for canonical and inverted orders, respectively. Like in the affirmative case above, pitch range in the

negative construction *ok ábí wúóyó kodí* ‘I will not talk to you’ is also expanded by being produced between 200Hz and 320Hz. Here the initial L-toned word *ok* ‘not’ is produced around 230Hz and the H-tones from the second syllable to the sixth syllable produced above 250Hz. The final H-toned syllable *-dî* in this clause falls back to 200Hz, as a downstepped !H. This can be due to the previous L-tone on *ko-* or due to the syllable being phrase-final. The second clause, in this case the subordinate clause, is compressed creating a flat F0 contour before dramatic lowering of the final L-toned syllable.

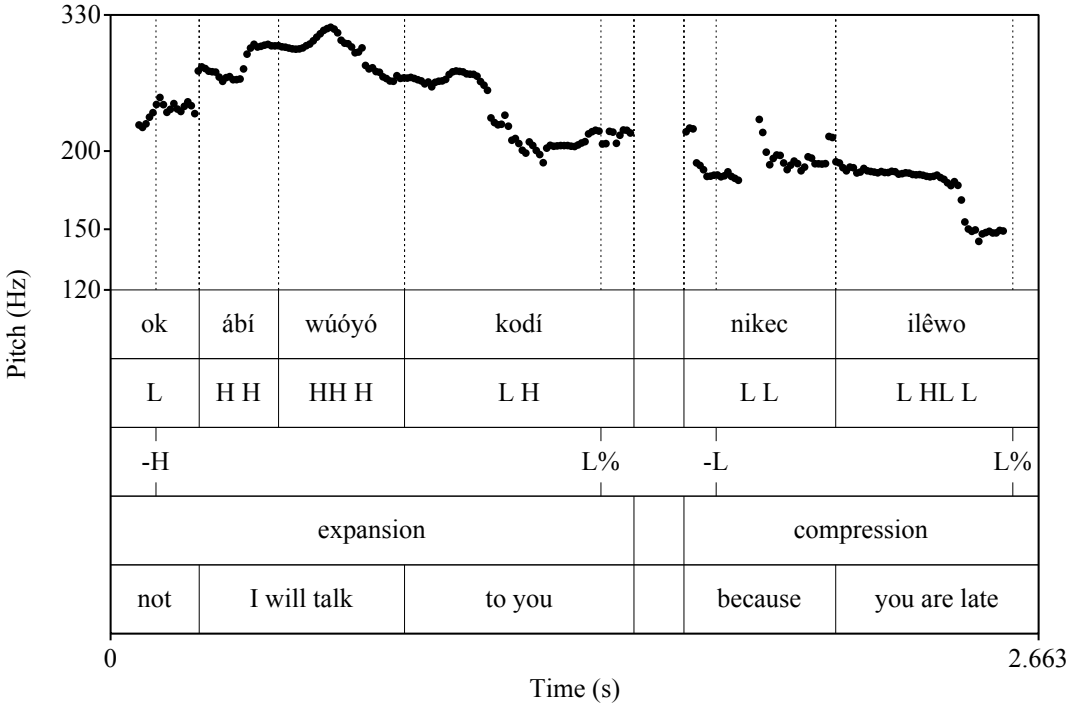


Figure 7.5 F0 contour of a complex sentence (103a) with a negative construction, showing PRE on the main clause and PRC on the adverbial clause

In Figure 7.6, the adverbial clause is fronted (in an inverted structure) but there is still pitch range expansion on the first clause and pitch compression on the second clause. Here, despite the embedded clause *nikec ilêwo* ‘because you are late’ having a sequence of L-tones in the

first three syllables it is produced above 200Hz while the H-tones on the second clause are produced below 200Hz. It is also observed that the pause in Figure 7.6 (when the subordinate clause precedes the main clause) is longer, with a difference of 24sec, than the one in Figure 7.5 (when the main clause precedes the subordinate clause).

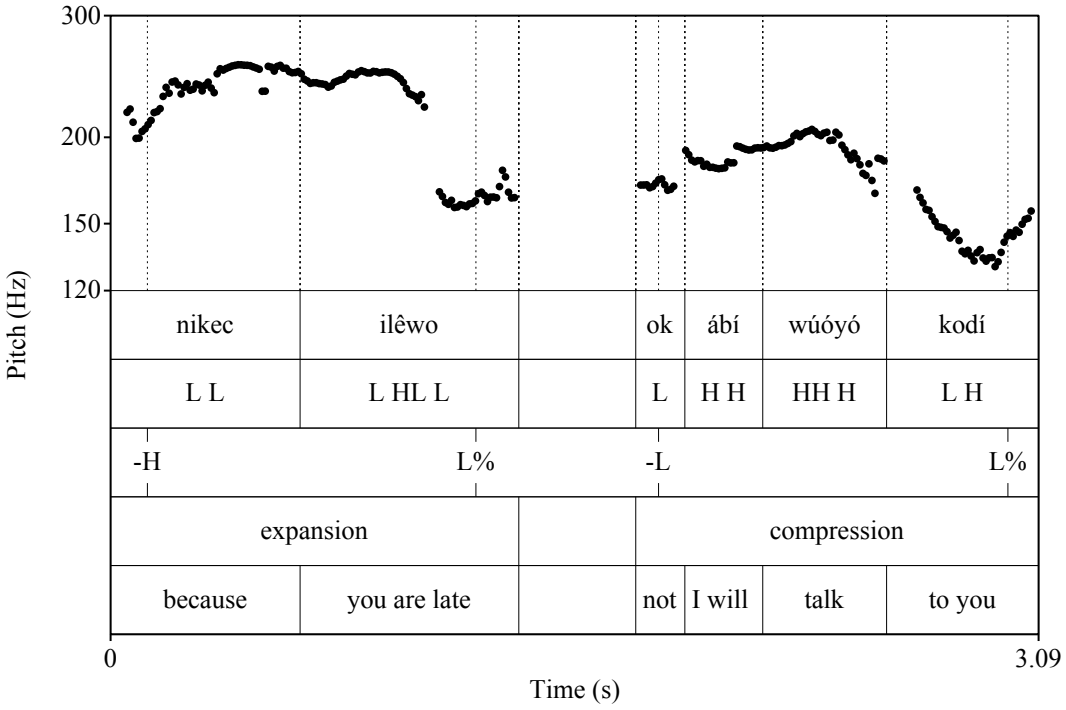


Figure 7.6 F0 contour of a complex sentence (103b) with a fronted adverbial clause followed by the main clause, showing PRE and PRC

Throughout Figures 7.3 to 7.6, we have seen that the major phonetic correlates of adverbial clauses are the global effects, PRE and PRC. The former affects the entire first clause while the latter affects the entire second clause regardless of whether it is an independent or dependent clause. In the previous discussion, we saw that relative clauses are recursive, being prosodically marked at their left- and right edges. Similarly, the adverbial clauses, as seen in the given examples, show recursive prosodic phrasing at the level of intonation phrases (IPs).

This is expressed by -H and -L boundaries at the left edges of the first and second clauses, respectively. The amount of register raising or lowering depends on the initial lexical tone. Generally, -H boundary is realized above 200Hz and -L boundary below 200Hz. Following the assumptions in Match Theory (Selkirk, 2011), this type of recursion shows that Luo prosodic constituents are isomorphic to syntactic constituents, as syntactic clauses in (104) match with intonation phrases in (105). However, having PRE constistently affecting the first clause and PRC the second, regardless of whether it is a dependent or independent clause shows that prosodic structure overrides syntactic structure. This is because the phonetic correlates which assign pitch to these syntactically different structures treat them as identical structures prosodically.

104. (a) [$_{CP}$ [$_{CP}$ né á-púró puoð-á] [$_{CP}$ ká né ogîk]]

[$_{CP}$ [$_{CP}$ I was cultivating my farm] [$_{CP}$ when he arrived]

(b) [$_{CP}$ [$_{CP}$ nikec i-lêwo] [$_{CP}$ ok á-bí wúóyó kod-í]]

[$_{CP}$ [$_{CP}$ because you are late] [$_{CP}$ I will not talk to you]]

105. (a) ($_{IP}$ ($_{IP}$ né á-púró puoð-á) ($_{IP}$ ká né ogîk))

($_{IP}$ ($_{IP}$ I was cultivating my farm) ($_{IP}$ when he arrived))

(b) ($_{IP}$ ($_{IP}$ nikec i-lêwo) ($_{IP}$ ok á-bí wúóyó kod-í))

($_{IP}$ ($_{IP}$ because you are late) ($_{IP}$ I will not talk to you))

7.1.3 Complementizer clauses

Lastly, for embedded (subordinate) clauses, another type which is investigated is the complementizer clause introduced by a complementizer *ni*.[16] While the relative and adverbial

[16] Luo has also copula *ni* as in *miyo ni gi kuon* (literally meaning, mother is with ugali) ‘mother has ugali’

clauses can be considered adjuncts as they modify the head noun or the verb or an adjective, the complementizer *ni* stands as an argument to the predicate. This means it can occur pre-verbally as the subject or post-verbally as the object of a sentence. Examples are given in (106) and (107).

106. ní ŋiθîndo ok o-cíèmo kaw-a gí wuoro
COMP children NEG PRF-eat take-1SG.OBJ by surprise
'That the children have not eaten surprises me.'

107. mama páró ní ŋiθîndo pókó o-cíè mo
mother think COMP children NEG PRF-eat
'Mother thinks that the children have not eaten.'

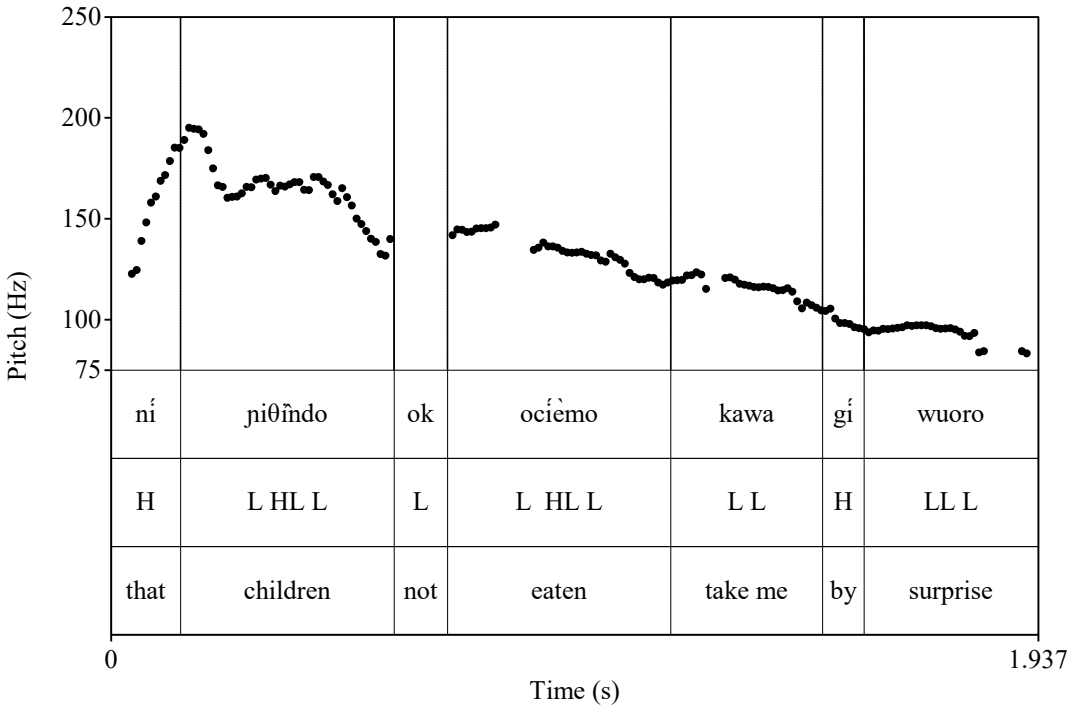


Figure 7.7 F0 contour of a complementizer clause (106) functioning as the subject of the clause, showing no edge effects

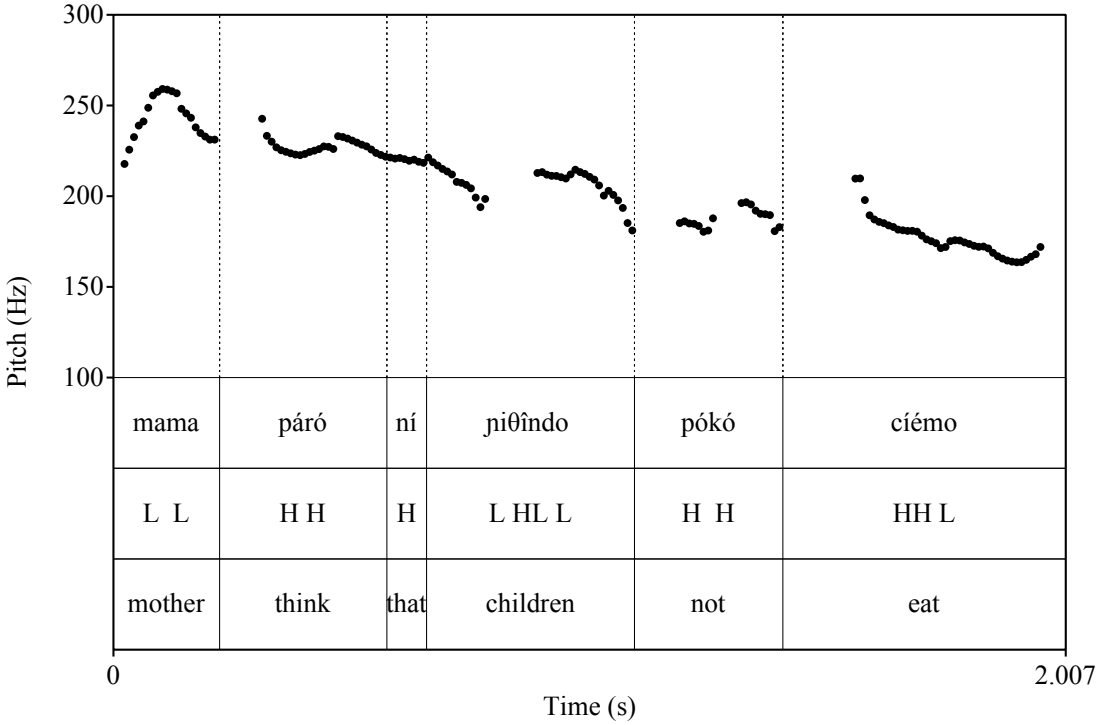


Figure 7.8 F0 contour of a complementizer clause (107) functioning as the object of the sentence, showing no edge effects

The pitch contours for (106) and (107) are illustrated in Figures 7.7 and 7.8 above, respectively. In these figures, it is clear that there are no left or right edge effects observed when the complementizer clause precedes or follows the verb. If they were present, at least in Figure 7.7, a L% boundary or a pause would be observed on the verb *o-cièmo* as an indication of a right edge effect, or else, there could be pitch reset on the first syllable of the verb *kawa*, which is the beginning of another syntactic constituent. Lack of vivid prosodic breaks is also observed in Figure 7.8 where the complementizer clause occurs after the verb as a sentential object. The contour patterning observed in complementizer *ni* is similar to that

of a simple declarative sentence discussed in chapter 5 section 5.1. In this section, it was argued that simple declaratives exhibit downtrend movements, as is also observed here. In this case, the Match Theory assumption by Selkirk (2011) that complex structures are intonationally recursive does not correctly predict the CP structures in (106) and (107). In their study of Zulu and Xhosa, Cheng and Downing (2007) also observed that there is no prosodic phrase break preceding sentential CPs. This suggests that lack of prosodic phrase breaks in complex clauses is not uncommon and that syntactic edges do not always trigger prosodic phrase breaks.

To sum up on subordinate clauses, in all the cases presented (except the complementizer *ni*) the syntactic clause (CP) is matched with a minor or major intonational phrase (ip/IP). The main phonetic correlates observed are pitch range expansion (PRE) and pitch range compression (PRC), which affect an entire ip or IP. This tendency can be attributed to anticipatory raising also known as pre-planning (Rialland, 2001; Gussenhoven 2004, Genzel, 2013). With pre-planning it is argued that speakers can anticipate the length or the complexity of a sentence and tend to begin higher.

7.2 Coordinated sentences

After examining the syntactic and prosodic structures of subordinated clauses, let us now turn to coordinated clauses. Syntactically, coordination or conjunction holds when two units are equivalent, meaning they have the same status or play the same role in a given syntactic context (Haspelmath, 2004; Fabricius-Hansen and Ramm, 2008). Thus, coordination, as opposed to subordination, is in a non-dependency relation.

Luo has the following coordinators: *gi* 'and', *kendo* 'and', *to* 'but/and'. Although they seem to overlap, they do play different roles. The first two coordinators *gi* and *kendo* express addition. *gi* is used to coordinate simple NPs as in (108a) or complex NPs as in (108b). *kendo* is used to coordinate clauses or VPs with an overt subject on the first conjunct and subject gap on the second conjunct, as in (109). In such cases, the subject of the first conjunct is also the subject of the second. On the other hand, *to* can have both addition and contrastive functions as shown in (110) and (111), respectively. According to Fabricius-Hansen and Ramm (2008), coordination can be asymmetric as in (108) or symmetric as in (110) and (111). The latter is found in 'standard' or 'canonical' coordination, which Fabricius-Hansen and Ramm argue is more of an exception than a norm.

108. (a) wuoro gi dayo
father CONJ grandmother
'Father and grandmother'

(b) dayo mo-tî gi ŋaθî matîn
grandmother REL.old CONJ child small
'An old grandmother and a small child'

109. wuoro tédó kuon kendo rúdó ŋuka
father PROG.cook ugali CONJ PROG.stir porridge
'Father is cooking ugali and he is stirring porridge.'

110. né wá-ði lim-é to n(é) ok wa-yud-é
PST SM1PL-go visit-OM3SG CONJ PST NEG SM1PL-find-OM3SG
'We went to visit him but we did not find him.'

111. wuoro twáŋó náŋgá to miyo tédó cíémo
father PROG.sew cloth CONJ mother PROG.cook food
'Father is sewing cloth and mother is cooking food.'

Given the purpose of this chapter, only clausal coordination will be examined. Since the coordinators *kendo* and *to* fit this purpose, as they connect constituents at the clausal level, their conjuncts will be prosodically examined. Figures 7.9 and 7.10 respectively illustrate the F0 contour of a sentence with *kendo* (109) and a sentence with *to* (111). First, it is noted that *kendo* and *to* are both phrased with the second conjunct. There is a pause between the first and the second conjunct which is followed by pitch reset at the left edge of the second conjunct. The reset in Figure 7.10 does not seem much higher given that the syllable *to* is L-toned and the final word in the preceding conjunct is lexically H-toned. As already argued previously, the amount of pitch reset depends on the lexical tones, as reset is superimposed on the lexical tone. Otherwise, pitch reset is evidenced by having an initial L-toned syllable of the second conjunct, in both figures, produced at relatively the same pitch range as the initial syllable of the first conjunct.

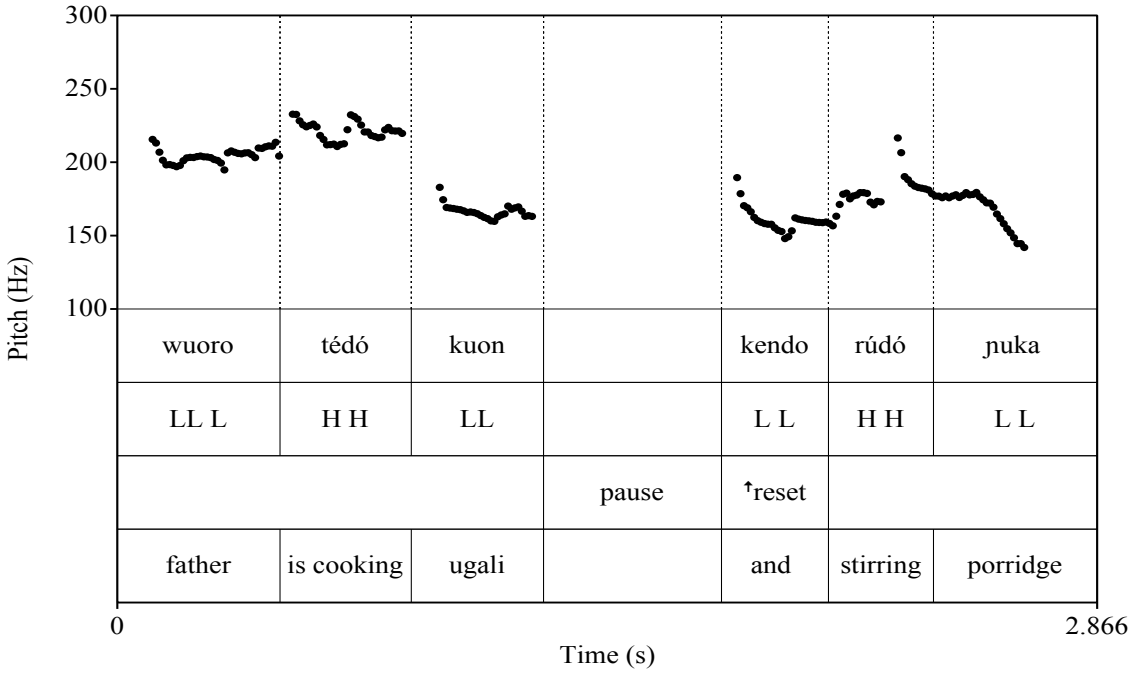


Figure 7.9 Coordinated complex sentence (109) showing a pause and pitch reset

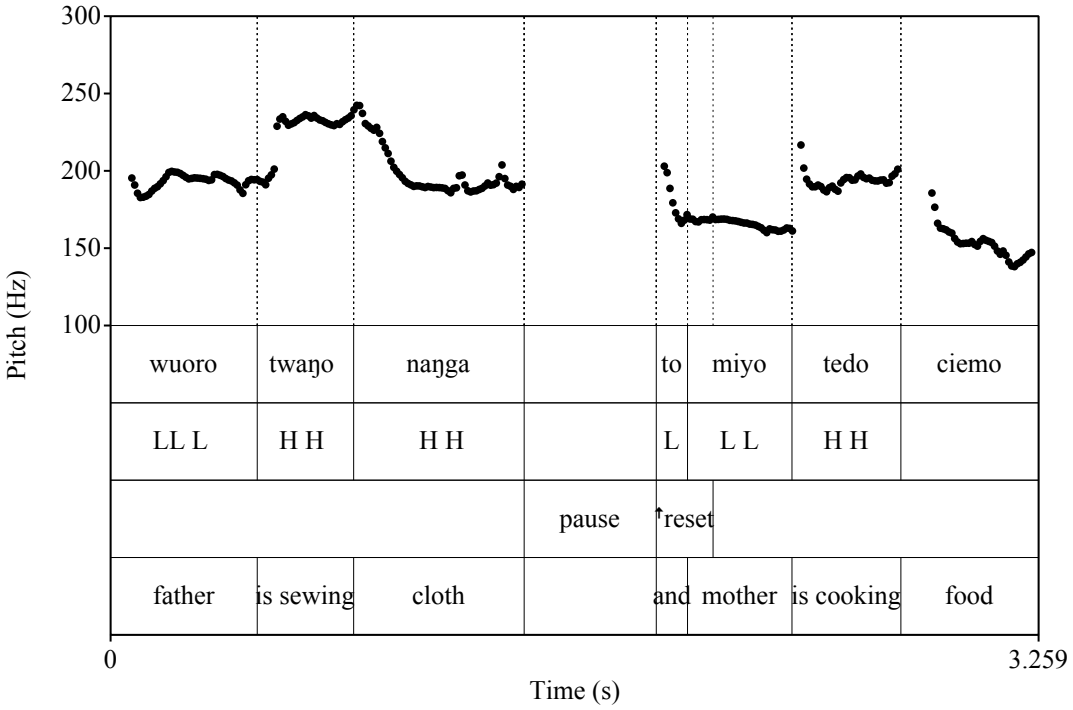


Figure 7.10 Coordinated complex sentence (111) showing a pause and pitch reset

Contrary to the subordinate clauses, the syntactic implication that coordinated clauses are independent is well reflected in their prosody. In subordinate clauses, we saw that the second clause is mainly compressed, while the second conjunct in coordinate clauses is not compressed. Having a prosodic boundary at the left edge of the second conjunct indicates that the coordinated structures are also recursive, being associated with intonation phrases (IP).

From the foregoing discussion of complex sentences, the following recursive structures in Table 7.1 are what emerge from all the sentence types considered.

Structure	Left edge	Right edge
Restrictive Relative Clause	-H (PRE); -L (PRC)	L%
Non-restrictive Relative Clause	Pitch reset	L%
Adverbial Clause	-H (PRE); -L(PRC)	L%
Complementizer Clause	None	None
Coordinated Clause	Pitch reset	L%

Table 7.1 Recursivity in complex sentences

This table shows that restrictive relative clauses have similar prosodic structure to that of adverbial clauses, with a -H boundary at the left edge of the first clause, as an indication of pitch range expansion and a -L boundary at the left edge of the following clause, as an indication of pitch range compression. Meanwhile, non-restrictive relative clauses and coordinated clauses are cued by pitch reset at the left edge. On the other hand,

complementizer clauses do not have left edge effects, which make them prosodically similar to declarative sentences.

7.3 Conclusion

In this chapter two main questions have been answered. The first one is: Are Luo complex sentences recursive? And if yes, how are they prosodically realized? The data presented in this chapter show that in Luo, relative clauses, adverbial clauses and coordinated clauses are recursive. And there is also a direct correspondence between syntactic edges and prosodic phrase breaks. The phonetic correlates for the embedded clauses are PRE and PRC, which are realized with -H and -L boundaries at the left edges, respectively. This is constistently observed in restrictive relative clauses and adverbial clauses. Additionally, there is pitch reset at the left edge depending on the tonal structure. This is observed in non-restrictive clauses and coordinated clauses. The right edges are also realized with boundary L% in all the embedded clauses. None of the embedded clauses are realized with a boundary H%. On the other hand, complementizer clauses formed with *ni* have no prosodic phrase breaks at the left or right edges.

Chapter Eight

Summary, Findings and Conclusions

8.0 Introduction

This thesis aimed to investigate the tonal and intonational structure of Luo. The study focused on determining the phonetic and phonological factors that contribute to the F0 contour of different sentences. Various questions, as outlined in chapter 1, are addressed. For convenience, these questions will be repeated here and their answers will be provided based on the analysis presented in chapters 3, 5, 6 and 7.

8.1 Summary of findings

In chapter 3, the main question addressed is how many tones does Luo have. In answering this question, different types of tones (i.e. lexical, grammatical, derived, underlying and surface) in Luo were analysed based on the F0 contour and the contexts in which they occur. The presented data show that Luo contrasts four lexical tones: High, Low, Falling and Rising. I considered H and L as the level tones in the language. However, the L-tone has a falling effect when it appears in isolated words. According to Yip (2002) this is due to a production effect especially when producing low-toned words. H-tone, on the other hand, is manifested at a sustained level. It is observed further that the distance between H and L is subtle. This, according to Maddieson (1978), is because the language contrasts only two-level tones. Tucker's (1994) claim that Luo has the so-called extra H-tone is disputed by the examples he provides. Data presented in section 3.1.1, shows that extra H-tone appears in derived phrasal contexts as a post lexical tone. Another argument was whether Luo has a downstepped !H (or

mid-tone as suggested in Tucker (1994)). Data in section 3.1.1 shows that downstepped !H is predictable and hence occurs as a derived tone only after another H-tone.

Regarding contour tones, it is observed that they can also occur as underlying lexical tones. This is contrary to Omondi's (1982) claim that contour tones in Luo are only derived. In Luo, as in other African languages, contour tones are treated as sequences of the level tones rather than units. Thus, a L and H can occur on a single syllable and yields a rising or falling contour. In addition, Tucker's (1994) claim, as shown in section 3.1.2, that Luo has several contour tones is phonologically implausible. Following the Autosegmental-metrical theory a language can have one contour as a phonological abstraction that captures the other configurations (i.e. high falling, extra high falling, long high falling, falling-rising and rising-falling).

Apart from the lexical tones, Luo also has grammatical tones, as discussed in section 3.1.3. Two grammatical functions are identified: aspectual distinction and word-class distinction. For aspectual distinction two tonal patterns, that is, H H H(L) and L HL L[17] are identified for progressive and perfective aspects, respectively. Following Odden's (1996) proposal I have argued that in the progressive aspect a H-tone is inserted on the first syllable of the verb and then spreads unboundedly to the rest of the verb. The tone does not spread to the final syllable when the verb is sentence-final. For the perfective aspect, a H-tone is inserted on the first syllable of the root and then surfaces as a falling contour. For the word class distinction, we have seen examples where adjectives can be distinguished from nouns using L and H tone, respectively.

---

[17] This pattern relies on the assumption that Luo verbs are mainly trisyllabic. For monosyllabic verbs the pattern will be H vs. HL, for disyllabic verbs HH vs. L HL.

For the case of derived tones, as presented in section 3.3, it is shown that three tonal processes: Tonal contraction, high tone spreading (HTS) and downstep produce different derivational tones. Based on Omondi's (1982) argument, tonal contraction derives a rising contour after attaching a floating H to the L-toned syllable. H-tone spreading is non-iterative and creates a falling contour when it spreads to a L-toned syllable or can surface as a H-tone when it spreads to a toneless syllable. Initial investigation shows that there are restrictions on the word/constituent boundaries that can be crossed when H-tone spreads. This aspect, however, needs further investigation and is reserved for future research. Lastly, downstep is seen to be the most prominent tonal process in Luo, as also the previous literature on the language suggests. Tucker and Creider (1975) and Tucker (1994) argued for four types of downstep: explicit, implicit, grammatical and built-in. All these fall under Yip's (2002) and Gussenhoven's (2004) automatic and non-automatic downstep. Automatic (explicit) downstep involves an overt L-tone that triggers downstep, while in non-automatic (implicit, grammatical and built-in) downstep there is no obvious L-tone trigger.

In chapter 5, the intonational analysis of Luo declaratives and questions is presented. The chapter addresses the following questions:

1. How are declaratives phonetically cued in Luo?
2. Are there any forms of register shifts as observed in other tonal languages?
3. Does Luo have final lowering? If yes, how is it grammaticalized? Do all four tones exhibit the final lowering effect or final lowering is restricted to some tones?

4. How do declaratives prosodically differ from questions?
5. Are question words intonationally marked?

In response to the first and second questions, controlled materials which comprise mixed and like-tones were analyzed and presented in section 5.1. Generally, both declaratives and questions exhibit a falling contour realized by register shift as in other tone terracing languages. This register shift is in the form of downstep or declination (i.e. downward movement). In Luo, there is no evidence of a floating L that triggers non-automatic downstep. What is observed in Luo is that non-automatic downstep is triggered by L% at the right edge of a phrase.

Another type of falling contour observed in Luo declaratives is declination. Unlike downstep, declination is a gradual fall in pitch that is attributed to phonetic rather than phonological effect. To measure its occurrence, Connell (2001) recommends the use of sentences with like-tones (all-high or all-low) rather than mixed tones. Following the quantitative test conducted, it is observed that an all-high tone sentence with six syllables (Figure 5.4) there is an average pitch drop of 15.25Hz, while an all-low tone sentence with ten syllables (Figure 5.5) there is an average drop of 3.58Hz. This indicates that longer sentences are realized with slower declination than shorter sentences (Genzel, 2013).

In response to the third question, it is observed that final lowering is also an important intonational cue in declaratives (as well as in questions). Final lowering is considered a dramatic fall at the right edge of a phrase. To measure the final lowering effect in Luo, I used

the exponential decay model proposed by Liberman and Pierrehumbert (1984). In this model, the peak values (in Hz) for both final and non-final L and H tones are measured to determine how they are related. The results show that final and non-final H-tones (in Figure 5.7) are asymptotic and thus lack abrupt final lowering. On the other hand, final L-tones, in Figure 5.8, are affected by final lowering. H-tones, phrase-finally are downstepped, generally exhibiting a flattened F0 contour. In this thesis, both incidences of final lowering and downstep are treated as final effects being associated with boundary L%. This is based on Gussenhoven's (2004) argument that down stepping a H-tone phrase-finally can be a form of grammaticalizing final lowering.

Concerning the prosodic differences between declaratives and questions, the results presented in section 5.2 show that both structures have falling F0 contours but questions are produced at a higher pitch register, commonly known as pitch range expansion (PRE). These results conform to what Creider (1978) found for Luo questions. PRE can affect the initial part of a question as in Figure 5.9 or the whole question as in Figure 5.10. This difference in the scope of PRE seems to be unsystematic as speakers differ in how they implement it. Phonologically, PRE is represented by a -H boundary at the left edge of an utterance. Alternatively, questions can have a rise-fall contour with reduced downstep as in Figure 5.11. However, the data on question prosody show that this type of prosody is non-native to Luo, as it was produced by two speakers when reading text but never in spontaneous speech. Note that spontaneous speech was used to complement the scripted sentences and they were referred to in some inconsistent cases such as when speakers produced different prosodic patterns. For that reason, I concluded that this is a language contact effect whereby speakers

mimic the HL question prosody found in Swahili, which is the dominant second language/lingua franca in the area. Finally, it is observed that question words are not prosodically marked.

In chapter 6, the intonational analysis of focus constructions, topics and dislocations is presented. All these reflect information structure which is about information packaging in a sentence. The general assumption that focus is prosodically prominent was taken as a starting point. A contribution to this discussion was made by investigating the phonetic manifestation of focused constituents in wh-questions (i.e for non-contrastive focus), contrastive constructions and cleft constructions. Their analyses are presented in sections 6.1, 6.2 and 6.3, respectively. The main question addressed here is how Luo marks focus. In the first case, wh-questions (i.e. content questions) were used. This is because question words in such questions usually seek some new information and thus considered to be inherently focused. The results, as presented in section 6.1, show that there is no prosodic marking on the focused constituent, but rather the whole sentence is realized at a higher pitch register, similar to that of wh-questions in chapter 5 (cf. Figures 5.10 and 6.1). The second case investigated was contrastive focus. For this case, data was obtained from parallel phrases, constructions with a set of alternatives and correction sequences. The results presented in section 6.2 show pitch raising mainly on the pre-focused constituent. In all the data used, the focused constituent is phrase-final, giving the impression that it may be affected by final effects. This makes it difficult to conclude on the scope of pitch raising and thus requires further investigation. Lastly, in the cleft constructions, it was noted that the cleft constituent (i.e. the XP) has pitch range expansion while the CODA is compressed. As in questions,

focused constructions, in general, have overall pitch range expansion marked by a -H boundary. But, the focused constituents (except cleft XPs) have no discrete prosodic marking.

Another question regarding the representation of information structure was whether dislocations are phrased together with or separate from the main clause. In section 6.4.1, it is argued that, structurally, both right-dislocated subjects and objects can occur at the right periphery in non-canonical order as long as their resumptive pronouns (i.e. referents) are marked on the verb. This means the right dislocation is morphosyntactically marked as it belongs to a separate CP. Prosodically, however, the F0 contours (see Figures 6.7- 6.8) show right dislocations phrased together with the host clause. This is realized by the lack of prosodic break between the host clause and the right-dislocated word.

Regarding the left-dislocated objects (i.e. topicalized objects), as discussed in section 6.4.2, it is observed that dislocated objects can occur with or without the resumptive pronouns on the verb. In both cases there is a prosodic break between the topic and the following clause. This break is realized by a boundary L% at the right edge of the dislocated/topicalized constituent and pitch reset at the left edge of the main clause. For the topicalized subjects, as discussed in section 6.4.3, it is necessary to have a resumptive pronoun on the verb which anaphorically refers to the lexical subjects for a structure to have a topicalized sense, as normally lexical subjects do not co-occur with their pronominal markers in Luo. Since topicalized subjects structurally behave like left-dislocated objects they are also prosodically set off from the main clause by a boundary L% at the right edge, but there is pitch reset at the left edge of the

main clause. It is thus concluded that Luo dislocations are asymmetric, whereby right dislocations are phrased together with the main clause while left dislocations are phrased separately from the main clause.

In chapter 7, an analysis of the intonation structure of complex sentences in Luo is given. The structures that were analysed include restrictive and non-restrictive relative clauses, adverbial clauses, complementizer clauses and coordinated clauses. The general assumption is that since complex sentences are made up of more than one clause, prosodically they are potentially recursive, creating multiple intonation structures (Selkirk, 1984; 2009). The prosodic analysis of embedded sentences, as presented in section 7.1, shows that relative clauses and adverbial clauses are recursive while complementizer clauses are not. As illustrated in Figure 7.1, the head noun and its modifier (i.e. the relative clause) are phrased together and produced at a higher pitch register, while the main clause is compressed. Similarly, for the adverbial clauses, as illustrated in Figures 7.3 – 7.6, a sentence-initial clause (i.e. adverbial clause in inverted order or main clause in canonical order) has expanded pitch while the sentence-final clause has compressed pitch, regardless of whether the preceding or the following clause is dependent or independent. Pitch expansion and pitch compression are respectively realized by a -H and a -L boundary tone at their left edges.

Non-restrictive relative clauses, like other embedded clauses, are also recursive. But unlike the restrictive relative, the non-restrictive relative is phrased separately from the head noun. Thus, the head noun, the relative clause and the main clause form different prosodic structures, which are realized by boundary L% and pitch reset at the right and left edges,

respectively. It is also observed that there are no pitch register differences between the recurring phrases in non-restrictive relatives (see Figure 7.2). As already pointed out above, complementizer clauses are non-recursive and lack prosodic breaks at the left or right edges. This makes a complementizer clause prosodically similar to a declarative sentence with an overall downward movement (i.e. declination or downstep).

Apart from the subordinate clauses, an investigation of the intonation structure of coordinated clauses was also conducted. According to Haspelmath (2004) and Fabricius-Hansen and Ramm (2008), coordination holds when two units are syntactically equivalent and have the same status. The analysis of such clauses is presented in chapter 7 section 7.2. Only clausal coordination was examined. Prosodically, the coordinated clause is also recursive, whereby the second conjunct is realized with pitch reset at the left edge, as illustrated in Figures 7.9 and 7.10.

The overall findings and conclusions of the thesis are therefore as follows:

- The tonal structure of Luo is composed of an equipollent system that contrasts H, L & Ø and includes two lexical contour tones HL and LH
- Downstep is the most significant process in Luo, contrasting automatic and non-automatic downstep, with no evidence of a floating L for non-automatic downstep
- Declination and final lowering both exist in the language with the rate of declination adjusting to the size of the sentence
- Final Lowering applies to both declaratives and questions
- Downstep is also a final effect triggered by a boundary L%

- Questions are produced with Pitch Range Expansion triggered by a left edge -H intonational tone. PRE can affect part or all of a sentence
- Questions words are prosodically unmarked
- There is no prominence on focused constituents but focus constructions are produced with a higher register
- Luo dislocations are asymmetrical, with right dislocations phrased with the main clause (even when constituents are pronominally marked), but left dislocations are phrased separately from the main clause
- Relative clauses (restrictive and non-restrictive), adverbial clauses and coordinated clauses are intonationally recursive, but complementizer clauses are not
- The intonational grammar of Luo uses the following set of intonational tones {L%, -H, -L, !H}

8.2 Future research

As with all work of this scale, there are a number of areas that warrant further investigation and where further work must be done. In particular, use of more spontaneous speech would be beneficial for any study of intonation; more data that elaborates more on the interaction of lexical tone and intonational tones; more analyses/understanding of the syntactic structure of sentences and contemporary analyses proposed for them, with a view to understanding what role intonation might have in this and specifically what implications the recursive structures in intonation plays in our understanding of sentence structure.

Furthermore, the findings herein focus on the Luo variety spoken in Tanzania. It is therefore important to venture into other varieties of Luo to see the extent to which these findings conform to (or not) to the other varieties of the language. These and others are issues to be pursued in future work.

References

Arvaniti, A. (2016). Autosegmental-Metrical Theory I [Slides] Aix Summer School on Prosody. University of Kent. 6 - 9 September 2016.

Arvaniti, A. & Baltazani, M. (2005). Intonational analysis and prosodic annotation of Greek spoken corpora. In S.A. Jun (ed) *Prosodic Typology and the Phonology of Intonation and Phrasing* (pp. 84 –117). New York: OUP.

Batibo, H. (2012). Does Kiswahili have diphthongs: interpreting foreign sounds in African languages. *South African Journal of African Languages,* 14, 180–186. 10.1080/02572117.1994.10587047.

Beckman, M. E., Hirschberg, J. & Shattuck-Hufnagel, S. (2005). The original ToBI system and the evolution of ToBI framework. In S.A. Jun (ed), *Prosodic Typology: The Phonology of Intonation and Phrasing* (pp. 9–54). New York: OUP.

Bell, J. (2005). *Doing your research project: A guide for first-time researchers in education, health and social science* (4th ed). Berkshire: Open University Press.

Bickmore, L. & Kula, N. C. (2013) Ternary spread and the OCP in Copperbelt Bemba. *Studies in African Linguistics,* 42(2), 101–132.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart & Winston.

Boersma, P. & Weenink, D. (2001). Praat, a system for doing phonetics by computer. *Glot International* 5:9/10, 341–345.

Bolinger, D. (1962). Binomials and pitch accent. *Lingua, 11*, 34–44.

Booij, G. (1995). *The phonology of Dutch.* Oxford: Oxford University Press.

Bruce, G. (1977). *Swedish word accents in sentence perspective*. Lund: Liber.

Cable, S. (2012). The optionality of movement and EPP in Dholuo. *Natural Language & Linguistic Theory, 30*(3), 651–697. Accessed from http://www.jstor.org/stable/41682600

Cahill, M. (2016). Kɔnni intonation. In L. J. Downing & A. Rialland (Eds.), *Intonation in African Tone Languages* (pp. 53–88). Berlin, Boston: De Gruyter.

Casali, R. (1997). Vowel elision in hiatus contexts: Which vowel goes? *Language, 73*(3), 493–533. doi:10.2307/415882

Cheng, L. & Downing, L. (2007). The prosody and syntax of Zulu relative clauses. *SOAS Working Papers Vol. 15* (pp. 51–63).

Cheng, L. & Downing, L. (2009). Where's the topic in Zulu? *Linguistic Review* 26, 207–238. 10.1515/tlir.2009.008.

Chomsky, N. & Halle, M. (1968) *The sound pattern of English*. New York: Harper & Row.

Clements, G. N. (1979). The description of terraced-level tone languages. *Language 55*, 536–558. Accessed from https://www.jstor.org/stable/413317 on 20-08-2019

Connell, B. A. (2001). Downdrift, downstep and declination. *Typology of African Prosodic Systems Workshop*. Germany: Bielefeld University.

Connell, B. A. (2016). Tone and intonation in Mambila. In L. J. Downing & A. Rialland (Eds.), *Intonation in African Tone Languages* (pp. 131–166). Berlin, Boston: De Gruyter.

Connell, B. A., Hayward, R. J. & Ashkaba, J. A. (2000). Observations on Kunama tone (Barka dialect). *Studies in African Linguistics 29* (1), 1–41.

Costa, J. & Kula, N. C. (2008). Focus at the interface: evidence from Romance and Bantu. In C. de Cat & K. Demuth (Eds), *The Bantu-Romance Connection: A Comparative Investigation of Verbal Agreement, DPs, and Information Structure* (pp. 293–322). Amsterdam: John Benjamins.

Creider, C. (1978). Intonation, tone groups and body motion in Luo conversation. *Anthropological Linguistics 20 (7),* 327–339. Accessed from http://www.jstor.org/stable/30027426

Cristofaro, S. (2003). *Subordination*. Oxford: Oxford University Press.

Daigle, J. (2008). *Tribes of Africa. 2nd edition*. Accessed from http://www.bellaonline.com/ebooks/ebook154.

Dimmendaal, G. (1995). The emergence of tense marking in the Nilotic-Bantu borderland as an instance of areal adaptation. In P. Zima (ed), *Time in Languages*, (pp. 29–43). Prague: Institute for Advanced Studies at Charles University and the Academy of Sciences of the Czech Republic.

Dimmendaal, G. (2002). Constraining disharmony in Nilotic: What does an optimal system look like? *Journal of African Languages and Linguistics*, 23. (pp. 153-181). 10.1515/jall.2002.010.

Doetjes, J., Rebuschi, G. & Rialland, A. (2004). Cleft Sentences. In F. Corblin & H. De Swart (Eds.), *Handbook of French semantics*, (pp. 529–552). Stanford, CA: CSLI Publications.

Downing, L. (2011). The prosody of dislocations in selected Bantu languages. *Lingua 121*, 772–786.

Downing, L. (2016). Tone and intonation in Chichewa and Tumbuka. In L. J. Downing & A. Rialland (Eds), *Intonation in African Tone Languages* (pp. 365–392). Berlin, Boston: De Gruyter.

Downing, L. & Rialland, A. (Eds). (2016). *Intonation in African tone languages*. Berlin, Boston: De Gruyter.

Downing, L. J., Mtenje, A., & Pompino-Marschall, B. (2004). Prosody and information structure in Chichewa. *ZAS Papers in Linguistics 37*, 167–187.

Fabricius-Hansen, C., & Ramm, W. (2008). Editor's introduction: Subordination and coordination from different perspectives. Accessed from https://www.researchgate.net/publication/275519268

Feldhausen, I. (2016). The relation between prosody and syntax: The case of different types of left-dislocations in Spanish. In M. Armstrong, N. Henriksen & M. del Mar Vanrell (Eds.) *Interdisciplinary Approaches to Intonational Grammar in Ibero-Romance*, (pp. 153–180). John Benjamins.

Genzel, S. (2013). Lexical and post-lexical tones in Akan. PhD dissertation. Institutional Repository of the University of Potsdam. URN urn:nbn:de:kobv:517-opus4-77969 http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-77969

Goldsmith, J. (1976). Autosegmental phonology. Ph.D. Dissertation, Cambridge, MA: MIT.

Gordon, M. K. (2005). Intonational phonology of Chickasaw. In S.A. Jun (ed) *Prosodic Typology and the Phonology of Intonation and Phrasing* (pp. 301–330). New York: Oxford University Press

Grice, M., Baumann, S. & Benzmüller, R. (2005). German intonation in Autosegmental-Metrical phonology. In S.A. Jun (ed) *Prosodic Typology and the Phonology of Intonation and Phrasing* (pp. 55–83). New York: OUP.

Grimes, B. F. (1996). *Ethnologue: Languages of the world* (13th ed, Dallas, Texas: Summer Institute of Linguistics). Retrieved from https://www.ethnologue.com/13/

Güldemann, T., Zerbian, S. & Zimmermann, M. (2015). Variation in information structure with special reference to Africa. *Annual Review of Linguistics 1*, 155–178. 10.1146/annurev-linguist-030514-125134.

Gussenhoven, C. (2002). Intonation and interpretation: Phonetics and phonology. *Proceedings of the First International Conference on Speech Prosody*. Aix-en-Provence, ProSig and Universit´e de Provence Laboratoire Parole et Language. 47–57.

Gussenhoven, C. (2002). Phonology of intonation: State-of-the-Article. *GLOT International 6* (Nos 9/10). 271-284.

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: CUP.

Gussenhoven, C. (2005). Transcription of Dutch intonation. In S.A. Jun (ed) *Prosodic Typology and the Phonology of Intonation and Phrasing* (pp. 118–145). New York: Oxford University Press.

Hale, M., Kissock, M. & Reiss, C. (2007). Microvariation, variation, and the features of universal grammar. *In Lingua*. Elsevier.

Hartmann, K. (2008). Focus and emphasis in tone and intonation languages. In A. Steube (ed.), *The Discourse-Potential of Underspecified Structures: Event Structures and Information Structure*, (pp. 389–412). Berlin, New York: De Gruyter.

Haspelmath, M. (2004). Coordinating constructions. *Typological Studies in Language*, 58. 10.1075/tsl.58.03has.

Hayes, B. (1989). The prosodic hierarchy in meter. In P. Kiparsky, G. Youmans (Eds.). *Phonetics and Phonology 1: Rhythm and Meter,* (pp. 201–260). San Diego: Academic Press.

Huang, C.T.J. (1985). On the autosegmental and metrical nature of tone-terracing. In D. L. Goyvaerts (ed). *African Linguistics,* (pp. 209–238). Amsterdam: John Benjamins.

Hyman, L. & Monaka, K. C. (2011). Tonal and non-tonal intonation in Shekgalagari. In S. Frota, G. Elordieta & P. Prieto (Eds), *Prosodic Categories: Production, Perception and Comprehension* (pp. 267–290). Springer Verlag.

Hyman, L. M. (2001). Privative tone in Bantu. In S. Kaji (ed), *Cross-linguistic Studies of Tonal Phenomena* (pp. 237–258). Tokyo: Institute for the Study of Languages and Cultures.

Hyman, L. M. (2009). How (not) to do phonological typology: The case of pitch-accent. *Language Science* 31(2–3)213–238.

Hyman, L. M. (2010). Amazonia and the typology of tone systems. In H. Avelino, M. Coler & L. Wetzels (Eds.), *The Phonetics and Phonology of Laryngeal Features in Native American Languages* (pp. 235–257). Leiden: Brill.

Johnson, R. B. & Christensen, L. (2014). *Educational research: Quantitative, qualitative, and mixed approaches*, 5th Ed. Thousand Oaks, CA. Sage Publications.

Jun, S. A. (2005). Korean intonational phonology and the prosodic transcription. In S.A. Jun (ed) *Prosodic Typology and the Phonology of Intonation and Phrasing* (pp. 2001-229). New York: Oxford University Press.

Jun, S. A. & Fletcher, J. (2014). Methodology of studying intonation: From data collection to data analysis. In S.A. Jun (ed) *Prosodic Typology II: The Phonology of Intonation and Phrasing*. Oxford: OUP.

Kanerva, J. M. (1990). *Focus and phrasing in Chichewa phonology*. New York: Garland.

Kihore, Y. M. (2005). Inflectional aspects in the Dholuo variety spoken in Tarime district, Tanzania'. *In LOT Occasional Papers in Linguistics 1*, 80.

Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica, 55* (34), 243–276.

Kügler, F. (2016). Embedded clauses and recursive prosodic phrasing in Akan. *Tonal Aspects of Language*, 24 –27.

Kügler, F. (2016). Tone and intonation in Akan. In L. J. Downing & A. Rialland (Eds). *Intonation in African Tone Languages*, 89–129. Berlin, Boston: De Gruyter.

Kula, N. C. & Bickmore, L. (2015) Phrasal Phonology in Copperbelt Bemba. *Phonology, 32* (1), 147–176.

Kula, N.C. & Hamann, S. (2016). Intonation in Bemba. In L. J. Downing & A. Rialland (Eds.), *Intonation in African Tone Languages* (pp. 321–364). Berlin, Boston: De Gruyter.

Ladd, R. D. (1996). *Intonational phonology*. Cambridge: CUP.

Ladd, R. D. (2008). *Intonational phonology (2nd Ed)*. Cambridge: CUP.

Languages of Tanzania Project. (2009). *Atlasi ya lugha za Tanzania*. University of Dar es Salaam.

Leben, W. R. (1973). Suprasegmental phonology. PhD Dissertation, MIT.

Liberman, M. Y. (1975). The intonational system of English. PhD Thesis, MIT.

Liberman, M. Y. & Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In M. Aronoff and R. T. Oehrle (eds), *Language and Sound Structure* (pp. 157–233). Cambridge, MA and London: MIT Press.

Maddieson, I. (1978). Universals of tone. In J. H. Greenberg (ed.), *Universals of Human Language, 2*: *Phonology*, 335–366. Stanford, CA, Stanford University Press.

Maddieson, I. (2013). Vowel quality inventories. In M. S. Dryer, & M. Haspelmath (Eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/2, Accessed on 2020-03-09.)

Makasso, E., Hamlaoui, F. & Lee, S. J. (2016). Aspects of the intonational phonology of Basaa. In L. J. Downing & A. Rialland (Eds), *Intonation in African Tone Languages* (pp. 167–194). Berlin, Boston: De Gruyter.

Nespor, M. & Vogel, I. (1986). *Prosodic phonology*. Dordrecht: Foris.

Odden, D. (1996). Tone: African languages. In Goldsmith, John. (Ed), *The Handbook of Phonological Theory*. Blackwell.

Odden, D. (2005). Introducing phonology. (Cambridge Introductions to Language and Linguistics.) Cambridge: Cambridge University Press.

Odden, D. & Bickmore, L. (2014). Melodic tone in Bantu: Overview. *Africana Linguistica,* 20, 3–13.

Odden, D. A. (1987). Predicting tone in Kikuria. In Odden, David A. (ed.), *Current approaches to African linguistics 4*, 311–326. Dordrecht/Cinnaminson: Foris.

Ohala, J. (1983). Cross-linguistic use of pitch: An ethological view. *Phonetica, 40*, 1–18.

O'Connor, J. D. & Arnold, G. F. (1973). *Intonation of colloquial English*, *2nd Edition*. London: Longman Group Ltd.

Ojal, G. (2015). The status of the vowel [A] in Dholuo: A phonological look at the vowel in terms of [ATR] value. *Journal Of Humanities and Social Science, 20* (5), 76–80.

Okoth-Okombo, D. (1997). *A functional grammar of Dholuo*. Köln: Köppe.

Okoth-Okombo, D. (1982). *Dholuo morphophonemics in a generative framework*. Berlin: Dietric Reimer Verlag.

Omondi, L. N. (1982). *The major syntactic structures of Dholuo*. Berlin: Dietrich Reimer Verlag.

Ott, D. & de Vries, M. (2016). Right-dislocation as deletion. *Natural Language and Linguistic Theory* 34, 641–690. https://doi.org/10.1007/s11049-015-9307-7

Palmer, H. (1922). *English intonation, with systematic exercises*. Cambridge: Heffer. Accessed from https://archive.org/details/englishintonatio00palmuoft

Patin, C. (2008) Focus and phrasing in Shingazidja. In M. Zygis & S. Fuchs (Eds.), *Papers in Phonetics and Phonology, ZAS papers in Linguistics, 49*, 167–189.

Patin, C. (2016). Tone and intonation in Shingazidja. In L. J. Downing & A. Rialland (Eds), *Intonation in African Tone Languages* (pp. 285–320). Berlin, Boston: De Gruyter.

Peng, S., Chan, M. K. M., Tseng, C., Huang, T., Lee, O. J. & Beckman, M. E. (2005). Towards a Pan-Mandarin system for prosodic transcription. In S.A. Jun (Ed) *Prosodic Typology and the Phonology of Intonation and Phrasing* (pp. 230–270). New York: OUP.

Pierrehumbert, J. (1980). The phonology and phonetics of English intonation. PhD Thesis. MIT.

Pierrehumbert, J. & Beckman, M. (1988). *Japanese tone structure*. Cambridge: MIT Press.

Rialland, A. (2001). Anticipatory raising in downstep realization: Evidence for preplanning in tone production. In S. Kaji (ed.), *Proceedings of the Symposium Cross Linguistic Studies of Tonal Phenomena: Tonogenesis, Japanese Accentology, and Other Topics* (pp. 301–321). Tokyo: Tokyo University of Foreign Studies.

Rialland, A. (2007). Question prosody: An African perspective. In T. Riad & C. Gussenhoven (Eds), *Tones and Tunes*: *Typological Studies in Word and Sentence Prosody* (pp. 35–62). Berlin: De Gruyter.

Rialland, A. & Aborobongui, M. E. (2016). How intonation interacts with tones in Embosi (Bantu C25), a two-tone language without downdrift. In L. J. Downing, & A. Rialland (Eds), *Intonation in African Tone Languages* (pp. 195–222). Berlin, Boston: De Gruyter.

Rialland, A. & Somé, A. (2011). Downstep and linguistic scaling in Dagara-Wulé. In John A. Goldsmith, E. Hume and L. Wetzels (Eds), *Tones and Features: Phonetic and Phonological Perspectives* (pp. 108–136). Berlin, Boston: De Gruyter.

Riedel, K. & Patin, C. (2011). Question structure and intonation in Fipa. *ZAS Papers in Linguistics, 55*, 141–160.

Rivero, M. (1980). On left-dislocation and topicalization in Spanish. *Linguistic Inquiry, 11*(2), 363–393. Retrieved from http://www.jstor.org/stable/4178165

Ross, J. R. (1967). Constraints on variables in syntax. Ph.D. Thesis. Massachusetts Institute of Technology.

Selkirk, E. O. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge. Mass: MIT Press.

Selkirk, E. O. (1986). On derived domains in sentence phonology. *Phonology Yearbook 3*, 371– 405.

Selkirk, E. O. (1996). The prosodic structure of function words. In J. Martin & K. Demuth (Eds.). *International Conference on Bootstrapping from Speech to Grammar in Early Acquisition.* Hillsdale, NJ: Lawrence Erlbaum.

Selkirk, E. O. (2000). The interaction of constraints on prosodic phrasing. In M. Horne, (Ed), *Prosody: Theory and Experiments,* 231–261. Kluwer.

Selkirk, E. O. (2002). Contrastive focus vs. presentational focus: Prosodic evidence from right node raising in English. *In SP-2002*, 643– 646.

Selkirk, E. O. (2009). On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure. In Gengo Kenkyuu 136, 35–73.

Selkirk, E. O. (2011). The syntax-phonology interface. In J. Goldsmith, J. Riggle, A. Y. (Eds). *The Handbook of Phonological Theory (2nd Ed).* Oxford: Blackwell Publishing.

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J. & Hirschberg, J. (1992). TOBI, a standard for labelling English prosody. In *ICSLP-1992*, 867– 870.

Stewart, J. M. (1965). The typology of the Twi tone system. *Bulletin of the Institute of African Studies,* Legon, 1, 1–27

Swenson, J. (2015). ATR quality in the Luo vowel system. *Canada Institute of Linguistics, EWP*. 1, 102–145 – via CanIL.

Truckenbrodt, H. (1995). Phonological phrases: Their relation to syntax and prominence. PhD Dissertation. MIT.

Truckenbrodt, H. (1999). On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry, 30*, 219–255.

Truckenbrodt, H. (2004). Final lowering in non-final position. *Journal of Phonetics, 32*(3) (pp. 313–348) Elsevier.

Truckenbrodt, H. (2007). The syntax-phonology interface. In P. de Lacy (Ed), *Handbook of Phonology* (pp. 435–456). Cambridge: Cambridge University Press.

Truckenbrodt, H. (2015). Intonation phrases and speech acts. In M. Kluck, D. Ott, & M. de Vries (Eds) , *Parenthesis and Ellipsis* (pp. 301–349). Berlin, Boston: De Gruyter.

Tucker, A. N. (1994). *A grammar of Kenya Luo (Dholuo).* Köln: Rüdiger Köppe Verlag.

Tucker, A. N. & Creider, C. (1975). Downdrift and downstep in Luo. *Proceedings of the Sixth Conference on African Linguistics.* OSU WPL20, 125–134.

Venditti, J. (1997). Japanese ToBI labelling guidelines. *OSU Working Papers in Linguistics, 50*, 127–162.

Venditti, J. (2005). The J_ToBI model of Japanese intonation. In S. A. Jun (Ed) *Prosodic Typology and the Phonology of Intonation and Phrasing* (pp. 172–200). New York: Oxford University Press.

Yip, M. (2002). *Tone.* Cambridge: Cambridge University Press.